ELUCIDATION OF STRUCTURAL AND FUNCTIONAL CHARACTERISTICS OF THE GUT MICROBIOME OF BEEF CATTLE UNDER WATER STRESS

By

HASITHA UPENDRA PREMATHILAKE

Bachelor of Science in Information Technology
Sri Lanka Institute of Information Technology
Colombo, Sri Lanka
2010

Bachelor of Science in Bioinformatics
University of Colombo
Colombo, Sri Lanka
2011

Master of Business Administration
Cardiff Metropolitan University
Cardiff, United Kingdom
2014

Submitted to the Graduate College of
The Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2021

ELUCIDATION OF STRUCTURAL AND FUNCTIONAL CHARACTERISTICS OF THE GUT MICROBIOME OF BEEF CATTLE UNDER WATER STRESS

Dissertation Approved:

Dr. Udaya DeSilva

_____

Dissertation Adviser

Dr. Chris Richards

_____

Dr. Darren Hagen

_____

Dr. Noha Youssef

_____

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my advisor Dr. Udaya DeSilva for the immense and continuous support and guidance given throughout my doctoral studies and related research, for his patience, motivation, and knowledge that made it possible for me to complete this journey. I could not have done it without you. I would like to thank my committee members Dr. Noha Youssef, Dr. Chris Richards and Dr. Darren Hagen for providing me with resources and assistance throughout the entire process of completing my dissertation research. I would also like to thank Dr. Adel Pezeshki for providing me with external research opportunities and Dr. Megan Rolf for assisting with various aspects of the doctoral research project.

I thank my current and former lab mates Dr. Xuwen Wienke, Chris Gaffney, Samantha Howe and Lulu Guo for all the help extended during lab work. A special thank also goes out to Rupi (beloved wife of Dr. DeSilva), for all the love, care and valuable advice provided, thus helping us through many difficult times. I would also like to thank all my friends in Stillwater, Oklahoma for all the help and companionship provided during our stay of four years at OSU Stillwater campus.

A special thank goes out to my beloved wife who has provided me with immense moral and emotional support throughout my post graduate research career and throughout the decade that we have been together. And also to my little daughter Sanuthi, for all the understanding and sacrifices made, tolerating times when I was pre-occupied with work. They are truly my pillars of strength and the cornerstones of my success.

I'm also thankful to Dr. Clint Rusk, Glynna Worley, Kathy Castre, all the other staff members and my fellow graduate students at Department of Animal and Food Science, for all the assistance provided in completing my research work. Last but not the least my loving farther and my late mother whose love, support and selfless sacrifices have made the person who I am today.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: HASITHA UPENDRA PREMATHILAKE
Date of Degree: JULY, 2021
Title of Study: ELUCIDATION OF STRUCTURAL AND FUNCTIONAL CHARACTERISTICS OF THE GUT MICROBIOME OF BEEF CATTLE UNDER WATER STRESS
Major Field: ANIMAL SCIENCE

Abstract:

The current trends in climate change and global warming are expected to have a profound effect on the cattle industry. Availability of good quality water in sufficient amounts is one aspect in cattle operations that can be adversely affected by such climate phenomenon. Water is an essential nutrient playing a pivotal role in maintaining critical physiological functions in cattle. The gastrointestinal microbial community of ruminants such as cattle are central to the digestion of plant material, production of volatile fatty acids, and the production of microbial crude protein essential in replenishing the nutrient requirement of the animal. Factors such as dietary composition, host genetics, production environment, age, sex etc. have been associated with significant changes in the gut microbiome of cattle. However, the effect of water restriction on the gut microbial dynamics of cattle is yet to be extensively studied. Hence, we used rumen and fecal samples from feedlot cattle collected during *ad libitum* water intake and at 50% water restriction) to reconstruct rumen and fecal microbial communities using 16S rRNA V4 gene amplicon sequencing and whole metagenome sequencing. The amplicon sequencing data summarized at genus level revealed significant differences ($p<0.05$) in the overall species composition of rumen and fecal microbiomes during water restriction. Genera such as *Methanobrevibacter, Rikenellaceae_RC9_gut_group* and *Prevotellaceae_UCG_003* showed significant differences in their relative abundance when subjected to water restriction. The fecal microbiome exhibited the most prominent changes due to water restriction where genera such as *Turicibacter, Clostridium_sensu_stricto_1, Christensenellaceae_R_7_group, Romboutsia,* and *Paeniclostridium* showed significant differences ($p<0.05$) in their relative abundance in comparison to *ad libitum* water intake. *Christensenellaceae_R_7_group, Paenoclostridum, Rombutsia, Clostridoides, Akkermansia* and *Lactobacillus* were identified as biomarkers in animals that performed significantly better ($p<0.05$) under water restricted conditions. Metagenome sequencing data summarized at species level showed a significant decrease ($p<0.05$) in the abundance *Ruminococcaceae bacterium P7, Methanosphaera sp BMS,* and *Methanobrevibacter millerae* in the fecal microbiome during water restriction. A multitude of biologically significant metabolic pathways in the rumen/fecal microbiome, pertaining to amino acid biosynthesis, methanogenesis, pyruvate fermentation etc. differed significantly ($p<0.05$) in pathway abundance during water restriction.

TABLE OF CONTENTS

Chapter Page

Chapter Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

Table Page

CHAPTER I

LITERATURE REVIEW

1.1 The ruminant animal

The first grasses appeared on the face of earth nearly 60 to 55 million years ago and started to evolve into the dominant form of flora. Another 5 million years down the evolutionary timeline, the first ruminants started to co-evolve as land (forest) dwelling animals in the form of small omnivores that has the ability to consume and effectively digest, the lignified plant material provided by the grasses (Kellogg, 2001; Weimer et al., 2009; Hackmann and Spain, 2010).

A ruminant can be defined as any member of the order Artiodactyla (belonging to class Mammalia) that possess a rumen, reticulum, omasum and abomasum. A total of 200 ruminant species that belongs to 6 families have been identified so far, out of which only 9 are domesticated. The wild ruminants are highly diverse in their phenotype, ranging from a bodyweight of less than 5kg to more than 800kg. The population of ruminants both wild and domesticated combined was a staggering 3.5 billion in 2010, out of which only an approximate 75 million are found in the wild. All domesticated ruminants belong to the family Bovidae, except the reindeer, whereas ruminants belonging to the other 5

families (Antilocapridae, Cervidae, Giraffidae, Moschidae and Tragulidae) remain in the wild (Hackmann and Spain, 2010).The typical diet of a wild ruminant can include natural forages such as grass, browse and fruits. Based on their differences in diet, ruminants can be divided into three main categories. The browsers, the grazers and the intermediate feeders. Grazers are animals who feed almost exclusively on grass, browsers on the other hand depend on a diet of twigs, tree leaves and herbs or forbes. The intermediate feeders feed on both grass and browse or change their preference based on the season. This difference in their dietary habits are reflected in the rumen morphology as well, even though it is still unclear whether the morphological adaptations actually represent the dietary differences, due to the lack of properly structured studies to address this particular aspect (Clauss et al., 2010).

The evolutionary history of ruminants suggest that they have predominantly been browsers throughout. Even today, the majority of them are classified as browsers and only a quarter are considered as grazers (Hackmann and Spain, 2010). Until recently, the popular consensus has been that both grazers and intermediate feeders evolved from browsers however, recent studies suggest that intermediate feeders were the more likely predecessor of both grazers and browsers (Clauss et al., 2010). In the context of domestic ruminants, cattle and sheep are well known grazers even though sheep are considered to be more selective than cattle. Goats on the other hand are mostly considered as intermediate feeders who rely on a mixed diet of grass, browse and forbs (Ferreira et al., 2017).

1.1.2 Importance of ruminants and rumen function

Ruminants are often considered to be superorganisms due to the prominent symbiotic relationship it maintains between the rumen microbiota and the cells of the host animal. This synergistic relationship comprises of interactions where the host animal provides food, heat and moisture to the rumen microbiota, in return the microorganisms provide the host with proteins and volatile fatty acids (VFAs) (Matthews et al., 2019).

The rumen is essentially a large pre-gastric fermentation chamber that is home to a complex ecosystem (comprising of anaerobic bacteria, methanogenic archaea, fungi, protozoa and phages) tailored to break down lignified plant material that are impossible for humans to digest. The domestication of ruminants possessing this unique ability to produce human consumables such as meat, milk, hides and wool from plant cellulose, has been the basis of animal agriculture for many generations (Weimer et al., 2009; Huws et al., 2018).

Milk in particular is a vital source of energy, enriched with vitamins, minerals and proteins, thus playing a pivotal role in food security. A 2012 world agriculture revision done by the Food and Agriculture Organization of the United Nations (FAO) estimated that, in order to meet the rising demand for animal proteins worldwide, a 76% increase in meat production and a 63% increase in milk production is required globally by 2050. In depth understanding of the rumen and its function, especially the rumen microbiome, is essential in devising new strategies to meet this demand. Studies on altering rumen function via the manipulation of the rumen microbiome through animal breeding and

dietary interventions, have gained momentum in the recent years with the ultimate goal of increasing milk and meat production and production efficiency, in order to cater to this growing demand (Huws et al., 2018).

Ruminants and the rumen function is also important as a potent source of methane, generated exclusively by the rumen archaea often referred to as the methanogens. Even though methanogens play an important role in maintaining the efficiency of the fermentation process, once the resulting methane is emitted to the environment via eructation, it becomes a source of atmospheric pollution (Matthews et al., 2019).

1.1.3 Structure and function of the bovine rumen

By 2010 the worldwide ruminant livestock population amounted to a staggering 3,612 million with sheep and cattle accounting for 95% of it. The sheer abundance of these domesticated ruminants makes ruminant fermentation the largest commercialized fermentation process in the world that breaks down lignocellulose, the most plentiful carbon polymer in the world (Weimer et al., 2009; FAO, 2013; Matthews et al., 2019).

At the heart of this process is the unique structure and function of the rumen, the largest of the four compartments that make up the complex stomach of cattle. The other three compartments are reticulum, omasum and abomasum. The rumen, reticulum and the omasum collectively forms the forestomach, where the brunt of the digestion of plant material takes place before it is passed on to abomasum, the smallest of the four compartments (Frandson et al., 2019).

The rumen, often called as the ruminoreticulum, due to being closely related to the reticulum both structurally and functionally, is divided into two compartments by muscular pillars. These are called dorsal and ventral sacs, the dorsal sac being the larger of the two sharing a dorsal space with the reticulum. The inner lining of the rumen is a mucous membrane made up of squamous epithelium that is both none-glandular and stratified. Both compartments of the rumen and the reticulum are lined with finger-like protrusions called papillae used for the absorption of nutrients such as VFAs (DePeters and George, 2014; Frandson et al., 2019).

The cyclic contractions of the ruminoreticulum facilitates the inoculation of plant material with microorganisms, the absorption of digestion end-products by papillae and the passing of digesta to the omasum. The gasses produced during the anaerobic fermentation process inside rumen gets passed into the esophagus, trachea and ultimately into the lunges which is then respired out to the atmosphere in a quiet process known as eructation (DePeters and George, 2014).

A healthy rumen pH is considered to be between 6 and 7. However, this may vary depending the on the type of diet the animal consumes. The buffering of the rumen pH largely happens via the enormous amount of saliva produced and swallowed by the animal, especially during cud chewing, and by the ammonia that gets produced during the anaerobic fermentation process. Maintaining a healthy rumen pH is important for its

proper function since it can affect rumen microbial activity and VFA production (Matthews et al., 2019).

1.2 US beef cattle industry

1.2.1 History

The first group of cattle imported into the US was by Christopher Columbus during his second trip in 1493. Records indicate that some Portuguese traders also imported beef cattle into North America in 1550s. However, the first large group of imported beef cattle were brought into the Jamestown Colony in 1611 by the British settlers. Early settlers from Spain, France, Netherlands, England and Sweden, all brought in cattle with them and pioneered the industry in states like Virginia, New York, Massachusetts, Delaware, and New Jersey. In the southwest, the first cattle are thought to have grazed around 1525, they were Spanish Longhorns brought in from Mexico, and by 1840 they were wide spread in states like Texas and New Mexico (Bowling, 1942; Wilson et al., 1965).

However, it would have been hard to fathom for the early European settlers how the beef industry in North America would grow in scope and size. A continuous paradigm shift has allowed the North American beef industry to expand itself from being a mere commodity to a more consumer driven business with increased value addition via science and technology innovations. The scope of modern US beef industry ranges from breeding, and raising beef cattle to processing, marketing, and merchandising of the end product to consumers (An Overview of the U.S. Beef Industry)

Records indicate a rapid increase in cattle numbers from early 1900s to 1975 where it peaked at 132 million. From 1975 onwards, there has been an overall decline in the annual cattle numbers in the US and despite the herd rebuilding process initiated in 2014, the 2015 cattle inventory indicated a total of 90 million and is expected hover around 89 to 93 million heads going forward (An Overview of the U.S. Beef Industry).

1.2.2 The economic impact of US beef cattle industry

By January 2020 the US cattle numbers stood at 94.4 million out of which 80 million was in beef production. According to USDA, in 2017, total sales pertaining to the U.S beef cattle industry stood at a staggering $77.2 billion, in 2019 it was $66.6 billion amounting to 18% of the total forecasted cash receipts for agricultural commodities in the United States. At the beginning of 2020, the beef production in the U.S. was predicted to be 27.5 billion pounds, essentially maintaining its 2019 numbers which stood at 27.15 billion with 33.6 million cattle being slaughtered (Derell, 2020; USDA ERS - Sector at a Glance, 2018).

For 2016, the annual cash receipts from agricultural products was expected to be an approximate $367.5 billion, with an expected contribution of $190 billion from livestock and livestock products and $74 billion (20.13%) of that coming from beef cattle. In terms of state wise contributions, 19 states had cash receipts greater than 1 billion USD, as per 2015 USDA agricultural statistics (An Overview of the U.S. Beef Industry).

Other than direct cash flows into the economy, many other industries that brings in billions to the U.S. economy such as finance, equipment, marketing etc., are supported by the beef cattle industry. One prominent example is the animal health products sector which recorded sales of about $2 billion USD in 2015. It is estimated that annual U.S beef industry income has a $3 to $5 multiplier effect on the overall U.S. economy (An Overview of the U.S. Beef Industry)

1.2.3 Current perspective

The United States is the largest consumer of beef, and is home to the world's largest fed-cattle industry, thus making it the most impactful livestock industry for the country's economy.

The U.S. beef cattle industry comprises of several different segments that are connected through beef production. These segments can be identified as;

- Seedstock segment
- Cow-calf segment
- Yearling stocker segment
- Feedlot segment
- Packing segment
- Purveyor segment
- Retail segment
- Consumer segment

The seedstock segment is comprised of specialized animal breeders who are responsible for the propagation of favorable genetics, with the intention of increasing profitability.

This is achieved via selective breeding and selling of animals, genetic information, semen and embryos to commercial cow-calf operations. The end product of the cow-calf segment is weaned calves which are then sold to yearling-stocker operations (to add more weight to weaned calves before they are sold to a feedlot) or feedlot operations. The weaned calves may remain in the feedlot for about 240 days, where they are fattened by feeding high concentrate diets, before being harvested. The harvesting of cattle usually takes place geographically in close proximity to feedlots. Beef packers, purveyors and retailers are responsible for the harvesting, processing and distribution of an approximate 24 billion pounds of beef annually. Beef packing is one of the most regulated businesses in the United States (An Overview of the U.S. Beef Industry; Drouillard, 2018).

The U.S. beef cattle industry has often been driven by technology. As a result, it has widely employed techniques such as supplementation of forage based diets to replenish the requirements of energy, minerals and vitamins for the animals. Furthermore, growth promoters are also widely used as feed additives or implants, with the exception of manufacturing programs that are branded as organic (non-hormone treated). Ionophore antibiotics such as monensin and lasalocid are also widely used in beef production in order to increase feed efficiency and prevent coccidiosis in cattle (Drouillard, 2018).

1.2.4 Future trends and challenges

In 2017, the U.S Food and Drug Administration (FDA) introduced new laws intended towards curtailing the use of medically important antibiotics without veterinary oversight. Drugs such as tylosin, oxytetracyline and chlortetracycline can now be only administered

under a prescription from a veterinarian and is no longer available “over the counter”. Important elements of the U.S beef supply chain, such as purveyors and major U.S retailers have clearly indicated their intention of moving towards “antibiotic-free” beef products and have introduced timelines for the procurement of such products. Major commercial beef producers have also implemented strategies that enable them to move towards an “antibiotic-free” production system without losing productivity. Therefore it is clearly evident that the U.S. beef cattle industry is moving towards an anti-antibiotic era, thus making the stakeholders to rely on scientific research and technological advancements to improve production practices in order to face off the inherent challenges this movement brings with it (Drouillard, 2018).

The increased use of probiotics can also be identified as an important future trend in the industry. It has been estimated that around 60% of the feedlot cattle are fed with probiotics at some point, citing reasons such as establishing normal GI tact functions and the inhibition of food-borne pathogens. A good example is the recent introduction of the bacteria *Megasphaera elsdenii* to the market, with reported benefits of improved cattle performance, prevention of ruminal acidosis, and decreased disease incidence in young cattle (Drouillard, 2018).

The tight regulation of antimicrobial drugs has given rise to another growing trend in the industry, the use of plant extracts as feed additives. Several plant extracts such as cinnamaldehyde, menthol, limonene and eugenol that are known to possess antimicrobial properties have been extensively studied, with notable impacts on the gut microbiota in

cattle. These studies have also suggested that the plant extracts can generate antibiotic resistance in bacteria, similar to conventional antibiotics. Other than plant extracts, heavy metal minerals such Zn are also increasingly used as feed additives to fight diseases such as respiratory illnesses and foot-rot in cattle (Drouillard, 2018).

In a global perspective, beef cattle operations are often flagged for their relatively high land occupancy and water usage. Studies suggest that about 8%-12% of energy in feed is lost due to methane production during the microbial fermentation process in the rumen, the same methane once released to the atmosphere, acts as a greenhouse gas. It has been estimated that the global livestock sector (of which beef and dairy cattle operations are major contributors) is responsible for about 14.5% of man-made greenhouse gas emissions. These, along with a range of other sustainability challenges, such as public health and animal welfare, will need to be addressed in a comprehensive manner to ensure the sustainability of beef supply chains, an imperative integral part in the drive for global food security (Gerber et al., 2015).

1.3 Global warming, climate change and its implications on the cattle industry

The last century and the years leading up to 2016 has been registered as the warmest on record where the modern human civilization has experienced the rise of average global temperature by $1.8F^0$ ($1.0C^0$). This period has given rise to record breaking weather extremes such as warmest years on record globally, and the current trend is expected to continue into the future as well. Extensive research has revealed the main culprit for this

global warming as anthropogenic greenhouse gas emissions world-wide (USGCRP, 2017).

Other than the rise in global atmospheric temperature, many other aspects of the global climate such as surface and atmospheric temperatures, snow and glacier cover, annual precipitation levels and the pH balance of sea water are also changing. The global sea level has increased about 7 to 8 inches since the beginning of the 20th century, and nearly a half of this increase has happened during the last 25 years. The rise of sea level is expected to continue into the foreseeable future with an increase of 1 to 4 feet predicted at the turn of the century, unless drastic measures are put in place to rectify the current global climate issues (USGCRP, 2017).

1.3.1 What's in store for North America

The climate changes pertaining to North America and the United States are well connected to global climate changes. As such the U.S. have also experienced a spike in its annual average temperature. For the period from 1895 to 2016 the increase of surface temperature in contiguous United State has been calculated to be somewhere between 1.2°–1.8°F (0.7°–1.0°C), which also supported by satellite image data that indicates a rapid increase of surface temperature from 1979 onwards. Furthermore, the annual mean temperature for the U.S is expected to climb in the future as well. For the period of 2021 to 2050 an increase of 2.5°F (1.4°C) is projected, much steeper rises are predicted for the late 21st century where lowest of the estimates would still place it at a 2.8°–7.3°F (1.6°–4.1°C) increase in the annual average temperature (Vose et al., 2017).

Temperature extremes like heat and cold waves have had significant changes as well. Compared to early 20th century the U.S. is now experiencing a lesser number of cold waves and a higher number of heat waves. Furthermore, compared to historical records, the temperatures of the extremely cold and hot days are predicted to increase while the total number of days that experience sub-zero temperatures are expected decrease in the U.S. Thus pointing towards a climate that is expected to significantly warm up in the years to come (Vose et al., 2017).

Even though national precipitation levels have seen an increase of 4% from 1901 and future precipitation events are also projected to increase, there are vast regional differences that needs to be considered. The West, South West and South East of the United States have experienced a decrease in annual precipitation levels while Northeast, Northern and Southern Plains and the Midwest have experienced an increase. During winter and spring the northern region of the U.S. is projected to experience an increase in precipitation levels while opposite is projected for southwestern United States (Easterling et al., 2017).

The increasing atmospheric temperatures trigger an increase in the capacity of water vapor it can hold, thus increasing the uptake of water vapor from land and water bodies. The decline of river flow, driven by the rising atmospheric and surface temperatures, have had a significant impact on the hydrologic paradigm of Northwest United States thus affecting river basins of Colorado and Rio Grande. Furthermore, recent studies have also pointed out river flow declines in the Rocky Mountains and the Missouri river basin

thus bringing in to light the expansion of hot droughts and arid conditions across the United States. Similar change has been observed in the Southwest as well, where declining river flows (since the latter half of the last century) are threatening the water supply of nearly a 40 million of its inhabitants and their livelihoods (Overpeck and Udall, 2020).

Even though other regions of the country are not expected to experience widespread aridification or long lasting droughts, overall sustained dry spells and flash droughts are expected to be the norm in the future. The most concerning factor is that this aridifcation of North America due to climate change is not expected to be reversible in the near future and the only hope lies in getting adapted to this "new normal" while trying to minimize anthropogenic greenhouse gas emissions that contribute to global warming (Overpeck and Udall, 2020).

1.3.2 Implications on beef cattle

By now it is a well-accepted norm that increased global warming and climate change is going to have extensive adverse effects on the cattle industry and livestock as a whole. The current trends justify increased concern about the thermal comfort of livestock including cattle, not only in tropical areas but also in areas that have a more temperate environment (Bernabucci, 2019).

Drought events has continued to adversely affect the U.S. beef cattle industry during the last decade. In 2011 and 2012, droughts affected more than 67% of the U.S. cattle

production and more than 70% of the overall livestock and crop production, and it is estimated that and approximate 14% of the United States can be affected by a drought at any given point in time (Countryman et al., 2016).

Heat stress is known to negatively affect the quality and quantity of milk production in dairy cattle, that can undermine the quality of origin cheeses, especially in the European region, thus degrading their reputation as niche products that boasts of excellence. Beef cattle are known to tolerate heat stress better due to their lower metabolic rates and body heat production. However, this tolerance comes at a cost where the increase in body temperature is compensated by homeostatic mechanisms and behavioral changes that may result in lower growth rates and reduced fertility (Bernabucci, 2019). Heat stress is also known to have a negative impact on the metabolism of glucose and lipids, decrease the adsorption of nutrients, alter liver function and create oxidative stress in cattle. Furthermore, a combination of reduced feed intake, reduced saliva production and a decrease in HCO_3^- content in saliva, in a heat stressed animal can make it more susceptible to rumen acidosis (Nardone et al., 2010).

Beef cattle that are fatter and with darker coat are particularly susceptible for heat stress due to extreme temperatures. It is suggested that the temperature range for beef cattle function without any adverse effect on daily weight gain is between $15C^0$ to $29C^0$. Beef cattle heat stress is also known to affect the quality of the meat produced, where an overall increase in pH and more dark cutting meat has been observed in such animals (Nardone et al., 2010).

Climate change, global warming and prolonged droughts can also have indirect effect on the cattle industry through decreased soil fertility, less availability of water, and increased circulation of pathogens other than direct effects such as heat stress. It is estimated that the yearly economic impact of heat stress can be as high as $897 million for dairy cattle and $369 million for beef cattle (Summer et al., 2019).

1.4 Beef cattle water intake

Even though overlooked in many of the nutrient requirement models, water is one of the most important nutrients required for the health and wellbeing of beef cattle. Water on average, accounts for 50% to 80% of an animal's body weight (BW) and plays an imperative role in feed digestion and maintaining other critical physiological functions. Hence, if an animal is deprived of sufficient amounts of the water, its feed intake may also drastically decrease, and even bodily functions may fail due to dehydration. Providing the animals with good quality water in sufficient amounts has always been a challenge for beef cattle production systems, especially during sustained droughts (Dyer et al., 2017; Wickramasinghe et al., 2019).

The minimum amount of water required by cattle to properly maintain its physiological functions, is determined by growth, water lost due to urinary excretions, sweat, feces and evaporation (through lungs or skin). Furthermore, the daily water requirement of an animal can be influenced by body weight, ambient temperature, stage of production, diet and the type of cattle. It is observed that the water requirement of beef cattle may double

when the ambient temperature increases from $50F^0$ to $95F^0$ and cattle feeding on hay and dry feeds would require a significantly higher water intake than cattle grazing on lush green pastures that can be 75% water (Rasby and Walz, 2011; Dyer et al., 2017).

Furthermore, it has been observed that introduction of water to new born calves since birth, rather than after 17 days of age (which is a common practice among farmers), may have asignificant impact in shaping their early gut microbiome. Calves that had access to water since birth had a significantly higher abundance of *Faecalibacterium prausnitzii*, and *Bifidobacterium breve* first 6 weeks of their life. These bacterial species are known to be associated with growth improvement in pre-weaned calves (Wickramasinghe et al., 2020).

1.5 The microbiome – An introduction

The concept of 'microbiome' first emerged as an extension of microbial ecology studies. It has quickly developed into a multidisciplinary field with applications in food & animal science, biomedical science, plant science and human medicine. The breakthroughs made in microbiome studies during the past two decades have widened our understanding about host-microbiome interactions, and have established the idea that eukaryotes in general are meta-organisms who are inseparable from their associated microbiomes, thus implying that the host and its microbiome is best considered as a single cohesive unit (Berg et al., 2020).

The definition of the word ‘microbiome’, has been debatable up to now since there is no clear consensus to which the experts in the field could agree upon. ‘Microbiota’ and ‘metagenome’ are also two other words that are often used interchangeably in microbiome studies. A most commonly cited definition for the microbiome by Lederberg done at the turn of the century, defines the microbiome as the microbial community that resides in a host or other environment while exhibiting commensal, pathogenic or symbiotic interactions with the host or the external environment. The total microbial community residing in a specified environment is best defined as the microbiota, whom together with its abiotic environment forms the microbiome. The word ‘metagenome’ is often used to describe the total genetic content arising from the microbiota residing in a specific environment (Marchesi and Ravel, 2015; Berg et al., 2020).

1.6 The rumen microbiome of cattle

The microbiota residing in the rumen, known as the rumen microbiome, comprises of a highly diverse population of microbes that plays a vital role in breaking down plant materials consumed by the animal into essential nutrients for the host (proteins, volatile fatty acids) and gases (CO_2, NH_3 and CH_4) that are emitted to the outside environment. This process is known as rumen fermentation, and the rumen microbiota plays unique, different and synergistic roles in this process (Fouts et al., 2012; Chaucheyras-Durand and Ossa, 2014).

The rumen microbiome is characterized by its high microbial density, diversity and complexity. It constitutes of a wide array of bacteria, fungi, protozoa and archaea,

bacteria being its most abundant component (McSweeney CS; Brulc et al., 2009; Mccann et al., 2014). The total biomass of the rumen microbiome is spread across three microenvironments; the liquid phase containing 25%, the solid phase containing 70% and rumen epithelial cells holding the remaining 5% (Ishler et al., 1996)

1.6.1 Structure and composition

Whole genomic DNA extraction, coupled with Next Generation Sequencing (NGS) to amplify and sequence bacterial marker genes or entire bacterial genomes, have become popular in rumen microbiome studies (Matthews et al., 2019). Both 16S rRNA amplicon sequencing and whole metagenome sequencing have been used to elucidate the structure and function of the rumen microbiome. By 2010, the Ribosomal Database Project (RDP) had curated a data base with data to identify an approximate 7000 bacterial species and 1500 archaeal species living in the rumen. Out of which 56% are Firmicutes, 31% are Bacteroidetes and 4% are Proteobacteria, making them the top 3 most abundant phyla. The same data at genus level reveals *Butyrivibrio, Ruminococcus, Acetivibrio, Succiniclasticum, Mogibacterium, Streptococcus* and *Pseudobutyrivibrio* as the predominant genera within Firmicutes. *Bacteroides* and *Prevotella* had the most database sequence read assignments within phylum Bacteroidetes (Chaucheyras-Durand and Ossa, 2014). In agreement with early seminal work on the rumen microbiome, recent large scale surveys such as The Global Rumen Census (GRC) project and the Hungate1000 project also reveals a rumen micobiome largely dominated by Firmicutes and Bacteroidetes, while the family *Lachnospiraceae* being described as the “largest single group” accounting for 32.3% of the total sequences obtained during the Hiungate1000

project (Seshadri et al., 2018). In a study comprised of 742 ruminant animals spanning across geographies, the GRC reports the existence of a core-microbiome, comprising of seven bacterial groups that account for 67.1% all bacterial sequences generated during the study. This core rumen microbiome consists of genera such as *Prevotella, Butyrivibrio, Ruminococcus,* and other taxa unclassified at genus level such as *Lachnospiraceae, Ruminococcaceae, Bacteroidales, and Clostridiales.* According to the GRC, a vast majority (70%) of the most abundant bacterial OTUs (Operational Taxonomic Units) are yet to be classified at genus level (Henderson et al., 2015).

The rumen archaea population is strictly anaerobic and is solely responsible for the production of methane. With a cell density of about 10^6 to 10^8 cells per ml of rumen liquid, archaeal members account only for less than 4% of the total microbial population residing in the rumen. *Methanobrevibacter, Methanomicrobium, Methanosphaera, Thermoplasma and Methanobacterium* are the five genera that account for almost 90% of the rumen archaeal population (Matthews et al., 2019). The methanogenic archaea in the rumen seem to be remarkably conserved across ruminant species and across geographies, according to the findings of the GRC project. *Methanobrevibacter gottschalkii* and *Methanobrevibacter ruminantium,* the two most abundant groups of methanogens accounting for almost 74% of the total archaea, were found in almost all of the samples. Together with "*Methanomassiliicoccaceae*-affiliated" and *Methanosphaera* sp., they accounted for 89.2% of all archaeal communities that were present in the samples (Henderson et al., 2015).

The rumen protozoan population was found to be less dense than the bacterial population at a cell density of 10^5 to 10^6 cells per gram of rumen contents, and was dominated by *Entodinium*, according to microscopic observations and real-time PCR (Skillman et al., 2006). Further assessments done using PCR primers specific to protozoans revealed three dominant protozoans at species level; *Entodinium caudatum*, *Epidinium caudatum*, and *Isotricha prostoma* (Sylvester et al., 2004). The GRC project found 12 protozoan genera in the samples collected throughout the globe. In agreement with historical studies, the *Entodnium* and *Epidinium* genera were found to be the most abundant, occurring in 90% of the samples and accounting for 54.7% of the total protozoal sequence data. In addition, genera such as *Enoploplastron* and *Ophryoscolex* were also found to be widely prevalent in samples collected from sheep and cattle (Henderson et al., 2015).

The rumen fungi population considered to be responsible for at least 10% of the microbial biomass in the rumen, has been described thus far through traditional culture-based methods, and the sequencing of the Internal Transcriber Region 1 (ITS1) of the rRNA gene. According to most recent literature, around 17 genera of anaerobic gut fungi (AGF) have been identified by culture based methods and confirmed by ITS1 sequencing (Guo et al., 2020; Hanafy et al., 2020). *Piromyces* was found to be the most abundant anaerobic fungal species in the rumen during a study conducted in 2010, using 30 different herbivore species, by sequencing the ITS1 of the rRNA gene in their rumen content. *Piromyces* accounted for 36% of the total ITS1 fungal sequences obtained during this study, while *Orpinomyces* and *Cyllamyces* were found to be the least abundant, represented by a mere 1.1% and 0.7% of the total sequences. Notably, a 38% of the total

sequences failed cluster alongside known fungal genera and formed separate clusters indicating the existence of novel anaerobic fungal taxa (Liggenstoffer et al., 2010; Krause et al., 2013).

Studies involving Transmission Electron Microscopy (TEM), molecular biological techniques and NGS have revealed a rumen virome that is dominated by bacteriophages. Early morphological studies conducted via TEM found a high abundance of phages that belong to the order Caudovirales. It contains viral families such as *Herelleviridae, Myoviridae, Podoviridae, Siphoviridae,* and *Ackermannviridae*. These viral particles have a common morphology of a polyhedral head and tubular tail. Further TM studies have also revealed the existence of viral particles that lacks a tubular tail, possibly belonging to viral families such as *Tectiviridae, Microviridae,* and *Corticoviridae* (Gilbert et al., 2020). Studies involving both molecular biological techniques such as pulse field gel electrophoresis and metagenomics studies involving NGS have further confirmed these findings and have also built on them. Unlike the rumen bacterial population that can be studied using both 16S amplicon sequencing and whole metagenome sequencing (WMS), the viral population can only be studied using WMS, due to the lack of universal primers that can amplify regions that are both variable and conserved enough to serve a similar purpose of the 16S rRNA gene (Gilbert et al., 2020). A recent study on rumen viral community of domestic caprids revealed the existence of double stranded DNA viral families such as *Siphoviridae, Myoviridae, Podoviridae, Mimiviridae, Microviridae, Poxviridae, Tectiviridae* and *Marseillevirus* (Namonyo et al., 2018). Another important study done on the bovine rumen virome revealed an absence of archaeal viruses and

rumen viral community abundant with viruses belonging to the families of *Myoviridae, Siphoviridae, Mimiviridae,* and *Podoviridae* (Anderson et al., 2017).

1.6.2 Functional dynamics

The consortia of microorganisms inhabiting the rumen and the hindgut of ruminants are responsible for the fermentation of feed components into short-chain volatile fatty acids. The ATP generated during this process facilitates microbial growth or the synthesis of microbial proteins that account for 60 to 85% of the amino acids found in the small intestine of the host (Hackmann and Firkins, 2015).

The anaerobic degradation of biomass inside the rumen can be divided into three main steps, namely; the breakdown of complex plant polysaccharides, the fermentation of carbohydrates into VFAs, and methanogenesis. Using rumen fluid as a proxy and utilizing a combined approach of metatranscriptomics and VFA profiling, a study conducted by Söllinger et al. revealed major microbial taxa and enzymes involved in these key steps. The degradation of plant polysaccharides was mainly achieved by four carbohydrate active enzymes (CAZymes) namely; cellulases, hemicellulases, starch-degrading enzymes, and oligosaccharide hydrolases. Transcripts coding for these enzymes were taxonomically assigned to *Prevotellaceae*, *Clostridiales*, *Fibrobacter*, *Ciliophora* (eukaryotic Ciliate), and *Neocallimastigaceae* (eukaryotic anaerobic fungi). The largest share of cellulase transcipts were produced by *Fibrobacter* and *Neocallimastigaceae.* Both *Clostridiales and Ciliophora* produced substantial amounts of cellulase and hemicellulase transcripts. *Prevotella* was found to express transcripts that

encode for hydrolases, starch degrading enzymes and hemicellulases, whereas *Firmicutes* were found to produce transcripts belonging to all four categories of CAZymes in approximately equal abundances. The expression of these CAZyme transcripts by several different microbial taxa suggests high functional redundancy in ruminal plant polysaccharide degradation (Söllinger et al., 2018).

The VFAs produced inside the rumen can largely be grouped into acetate, propionate and butyrate. Transcripts associated with acetate production (via pyruvate, acetyl-CoA, and acetyl-CoA & acetyl-P pathways) were mainly produced by *Prevotellaceae* and *Firmicutes* with *Prevotellaceae* accounting for the large majority. The production of acetate in the rumen happens via two main pathways, namely, the succinate pathway and the acrylate pathways. Transcripts encoding for both of these pathways were produced by *Prevotellaceae and Clostridiales,* and again *Prevotellaceae* contributed for a large majority. In contrast, transcripts involved in butyrate production through butyrate kinase pathway and butyryl-CoA:acetyl-CoA transferase pathways were only associated with *Firmicutes (*i.e *Clostridiales* and *Negativicutes).* Notable animal to animal variation was also observed during this study, stemming from the differences in the relative abundance of contributing taxa (Söllinger et al., 2018).

Methanogenesis by methanogenic archaea in rumen, can happen via three pathways namely; hydrogenotrophic pathway, methylotrophic pathway and acetoclastic pathway, with hydrogenotrophic pathway powered by *Methanobrevibacter,* being the most common. Methanogenesis via methylotrophic and acetocalstic pathways are not common

due to the low abundance of methanogens such *as Methanomassiliicoccaceae, Methanosarcinales and Methanosphaera* (Huws et al., 2018).

1.6.3 Interaction with external factors

Host genetics

Out of the many factors that are known to affect the rumen microbiome, the effect of host genetics maybe one of the least studied. Some recent studies published within the last five years have been able to shed some light on this. A study published in 2019 by Fan et al. (Fan et al., 2020) tried to investigate the effect of host genetics on the gut microbiome by creating a "graduated spectrum" of host genetic variation by breading 228 calves that represented a linear variation in genetic composition from Angus (100% *Bos taurus*) to Brahman (100% *Bos indicus*) and analyzing their fecal microbiome composition and function. They uncovered that host genetics, predominantly of paternal origin, had a significant effect on the gut microbiome of pre-weaning calves. They also found that single nucleotide polymorphisms (SNPs) in the genes that code for mucin (imperative for the wellbeing of gut mucosae) in the host genome had a significant relationship with gut microbiota responsible for mucin degradation. However, this study was unable to draw any conclusion about the effect of host genetics on the rumen microbiome since, at the time of sample collection, calves did not have a fully developed rumen (Fan et al., 2020).

The study by Roehe et al. (Roehe et al., 2016) describes how bovine host genetics was found to have a significant impact on the methanogen microbiota in the rumen microbiome. The genetically different sire progeny groups used for this study, when

ranked according to their methane emissions, demonstrated a significant correlation with the relative abundance of methanogen archaea, thus implying that at least the methane-producing microbiota in the rumen is under considerable host genetic control. Furthermore, this also suggests that producing low methane-emitting animals, based on the host genetic control of methanogens in the rumen microbial community, is a viable option (Roehe et al., 2016).

In 2019, a study based on 709 beef cattle comprising of Angus, Charolais and Kinsella hybrids, revealed that rumen microbiota of this commercial cohort differed mainly on diet, breed and sex with additive host genetic control having a considerable impact on its segregation. Around 59 microbial taxa were found to be moderately heritable and the majority of them being keystone members of the co-occurrence network of rumen microbiota. They were also found to be significantly correlated with host traits such as volatile fatty acid (VFA) concentration and feed efficiency (Li et al., 2019).

Ambient temperature

The effect of ambient temperature, relative humidity and antimicrobial compounds in diet are some of the external factors arising from the animal environment that can affect the rumen microbiome (Tajima et al., 2007; Romero-Pérez et al., 2011). A study conducted by Romero-Perez et al. in 2011 (Fan et al., 2020) observed that ambient temperatures largely had an indirect effect on rumen microbiome. However, proportional differences based on ambient temperature was detected in some bacterial taxa at both phylum (*Firmicutes*) and order (*Lactobacillales*) levels (Fan et al., 2020).

In another study conducted in 2006 by Tajima et al. (Tajima et al., 2007) that consisted of 3 sets of experiments, the animals in the first experiment did not exhibit significant difference in their rumen microbiome based on temperature differences that ranged from $20C^0$ to $33C^0$. However, when coupled with 80% humidity, they were able to detect significant differences in the gut microbiome (Tajima et al., 2007).

Heat stress

Short-term or prolonged acute heat stress has been demonstrated to have detrimental effect on the rumen microbiome. A recent study conducted by Baek et al. (Baek et al., 2020) reported that Hanwoo steers under 6 consecutive days of heat stress (housed in chambers maintaining 60% humidity and 35^0C ambient temperature) had an increase in the abundance of *Ruminobacter*, Lactobacillaceae and *Prevotella*, while species belonging to genus *Ruminococcacea*e showed a decrease in their abundance. However, rumen archaeal populations representing the bulk of rumen methanogens did not exhibit a significant impact from heat stress (Baek et al., 2020).

A separate study conducted using lactating Holstein dairy cows (n=18) showed similar results with no significant impact on the overall diversity of the rumen microbiome. A comparison of control and heat stress groups of this study revealed increased relative abundances of *Streptococcus*, *Enterobacteriaceae*, *Ruminobacter* and *Treponema*, and a decrease in the abundance of *Acetobacter* in the rumen of heat-stressed animals (Zhao et al., 2019).

Diet consumed and consumption level

Since the dietary composition of beef cattle vary significantly during their lifespan, the rumen and its microbiota needs to adjust accordingly as well, resulting in changes in the structure and function of the rumen microbial community. A study published in 2010 by Pitta et al. (Pitta et al., 2010) evaluated the changes in rumen microbiota during the transition from Bermuda grass to wheat pastures in beef cattle operations. The relative abundance of *Prevotella* was observed to decrease while the contrary happened to the abundance of *Rikenella* during the transition, along with an increase in the abundance of *Succiniclasticum*. A decrease in overall community diversity of the rumen microbiome was also observed as an effect of transitioning into a high energy diet (Pitta et al., 2010).

An increase in the abundance of *Olsenella, Atopobium, Lactobacillus, Desulfocurvus*, and *Fervidicola* in the rumen microbiome was connected to high-grain diets, while for cattle grazing on wheat pastures, an increase in the ratio of Bacteroidetes:Firmicutes was observed usually associated with increased biofilm formation that may also result in frothy bloat (Loor et al., 2016). Bacterial phyla such as Bacteroidetes and Firmicutes are known to affect residual feed intake and efficiency in Holstein dairy cows. It has been reported that animals that possess high feed efficiency and RFI, have a significantly high abundance of Bacteroidetes while an insignificant but nevertheless comparatively low abundance of Firmicutes was also noted (Delgado et al., 2019).

1.7 Next generation DNA sequencing (NGS)

1.7.1 Sanger sequencing

The dideoxy chain termination method (Sanger sequencing) developed by Fredrick Sanger in late 1970s laid the foundation for the DNA sequencing technologies as we know it today. This method relies on dideoxynucleotides (ddNTPs) that lack the 3' hydroxyl group, thus enabling chain termination during a DNA extension reaction (PCR). Incorporation of a low concentration of these ddNTPs into a PCR reaction triggers early termination and generates DNA strands of each possible length, as these ddNTPs get randomly incorporated into the synthesized DNA stand. The sequence of a given DNA template can be observed running four such parallel reaction of each kind of radiolabeled ddNTPs (A,G,T,C), and then running them in parallel on a polyacrylamide gel (Heather and Chain, 2016).

By 1980s the scientists at California Institute of Technology were able to automate Sanger's original method, which was then commercialized by the manufacturing of commercial First Generation sequencing machines by Applied Biosystems. In these machines the polyacrylamide gel was replaced with capillaries and radiolabeled ddNTPs were replaced with fluorescent labeled ones that generates an electropherogram which can be used for base calling, enabling its widespread use. Sanger sequencing is best suited for the sequencing of fragments that are around 500bp-1000bp in length. First Generation sequencing machines played a pivotal role in the Human Genome Project, and are still in use today, where sequencing of larger DNA fragments is required but low

throughput and high cost sequencing is acceptable (Sequencing 101: The Evolution of DNA Sequencing Tools - PacBio).

1.7.2 Second generation sequencing

Alongside the large scale and commercial use of automated Sanger (first generation) sequencing technologies, the very first massively parallel, second generation sequencing technology was developed by 454 Life Sciences, based on a technique known as pyrosequencing, originally developed by Pal Nyren and colleagues. This parallelization enabled high-throughput sequencing (HTS) that drastically increased the yield of DNA sequencing projects making it possible to sequence the genome of a single human in record time, overtaking similar efforts made at the Craig Venter institute, that was using first generation sequencing technologies (Heather and Chain, 2016).

Overall, second generation technologies can be divided into two broad categories; optical sequencing and non-optical/post-light sequencing. Optical sequencing includes 454-pyrosequencing based sequencers and Solexa/Illumina sequencing (MISeq, HiSeq and NextSeq sequencers). Whereas non-optical sequencing include semiconductor based sequencers such as Ion Torrent and S5 sequencers from ThermoFisher (Singh, 2017).

Illumina sequencing or reversible terminator-based sequencing uses distinctly labeled nucleotides in a chain termination, sequence by synthesis (SBS) approach to identify nucleotides in the template DNA strand. The template DNA strands are made of clonally amplified input DNA. Once the labeled nucleotides are bound (entire flow cell is flooded

with labeled dNTPs and DNA polymerase), the template strands are imaged and fluorescence is recorded. However, in contrast to Sanger sequencing, these chain terminations are reversible hence the bound dNTPs are washed away and the flow cell is prepped for the next “cycle” where the subsequent nucleotide of same clonal DNA strand will be sequenced. This is the most widely used second generation sequencing technology to-date, and forms the basis for MISeq (medium throughput, table-top), HiSeq (high throughput, commercial) and NextSeq (medium-high throughput) sequencers from Illumina (Singh, 2017).

1.7.3 Third Generation sequencing

One of the main limitations in second generation sequencing is that the genomic input DNA needs to be fragments in relatively short sequences. Furthermore, the clonal amplification of fragmented genomic DNA introduces PCR bias and makes it difficult to quantitatively compare end-results (this specifically affects microbiome projects). Hence, Single Molecule Real-time (SMRT) sequencing, or third generation sequencing has been looked upon as a promising solution for these issues (Heather and Chain, 2016).

PacBio and Oxford Nanopore Technologies (ONT) currently offer SMRT solutions, however, PacBio is considered to be the most widely used. It uses adapter ligation to double stranded DNA to create ‘circular’ DNA that is called SMRTbell library. DNA polymerase and labeled nucleotides are added to the library and placed on the SMRT cell that contains nanostructure wells called zero-mode waveguides (ZMWs). Each ZMW will hold a single double stranded DNA ligated with the adapter and DNA polymerase

attached. When DNA polymerase adds labeled nucleotides to the template strand, fluorescence is detected and recorded. ONT, on the other hand, utilizes a completely different approach for single molecule sequencing, where the DNA molecule is unzipped and the template strand is threaded through by the Motor Protein, a protein nanopore fixed to a membrane where voltage is applied so that there is ionic current flow through pore. When template strand goes through the pore, this ionic current is disrupted, and these disruptions can be used to identify the nucleotide that passes through the pore (Lu et al., 2016).

A major drawback of ONT is its high error rate, and the requirement of specialized sequence assembly software to assemble the long DNA sequences obtained via third generation sequencers. However, the advances in *denovo* sequencing assembly techniques have paved the way for the combinatorial usage of Illumina and ONT/PacBio technologies where long read assembly and short read error correction can be used to assemble complete bacterial genomes in microbiome studies (Heikema et al., 2020; Moss et al., 2020).

1.8 16S rRNA gene sequencing for microbiome studies

1.8.1 The 16S rRNA gene

The 16S, 5S and 23S rRNAs present in the ribosomes of bacteria and archaea are coded by the ribosomal RNA operon (rrn), available in multiple copies usually of 1-15 in bacteria and 1-4 in archaea (Stoddard et al., 2015). The 16S rRNA molecule that forms the Small Subunit (SSU) of the prokaryotic ribosome contains interleaving conserved and

hypervariable regions that are also different in their rate of molecular evolution, thus enabling their utilization as molecular biomarkers to derive both ancient (domains) and recent (genera) taxonomic lineages. The secondary structure of the 16S rRNA that consists of nearly 50 helices is highly conserved as well, making them important structural features that establish positional homology in phylogenetic analyses (Yarza et al., 2014).

Similarly, this interspersed hypervariable and conserved structure is preserved in the gene for 16S rRNA as well. The hypervariable regions are unique to different bacterial species or genera and the conserved regions make it possible to design universal primers that enable PCR amplification and sequencing of the variable regions. However, the 9 hypervariable regions all differ in their degree of sequence diversity. Hence, the efficacy of each hypervariable region in distinguishing different bacterial taxa vary, and no single hypervariable region has the capability to classify all bacteria correctly (Chakravorty et al., 2007; Ames et al., 2017).

1.8.2 Taxonomic profiling of the microbiome through 16S rRNA gene amplicon sequencing

The availability of a plethora of hypervariable regions whilst being highly conserved and abundantly available across all archaea and bacteria have made the 16S ribosomal RNA (rRNA) gene the preferred molecular phylogenetic marker for many taxonomic studies from as early as 1977 (Kim and Chun, 2014).

With the advent of NGS, 16S rRNA gene-based taxonomic profiling has become a popular and cost-effective technique in studying the composition of a microbial community, that does not rely on culturing. Its ability to determine the relative abundance of bacterial taxa in a sample, increased efficiency that enables the sequencing of a multitude of samples simultaneously, improved turnaround times, and its ability to be utilized for the surveillance of pathogenic bacteria in a clinical setting (through amplicon sequencing of selected variable regions) have further contributed to increased popularity (Gupta et al., 2019).

Existing literature suggests two different approaches when utilizing 16S rRNA sequences for microbiome studies; one targeting only a specific region comprising of a single or a few adjacent hypervariable regions, and the other being the amplification and sequencing of the entire 16S rRNA gene. The former being used for taxonomic identification at genus level or above, and the latter being utilized for more precise species and strain level identification (Kim and Chun, 2014; Johnson et al., 2019)

Targeting a single hypervariable region (e.g. the V4 region) has long been a norm in 16S studies. However, due to the limitations mentioned above and due to the huge progress made in next generation sequencing technologies, current and future studies have an increased tendency to be based on sequencing the whole 16S gene. Optimized circular consensus sequencing techniques coupled with state-of-the-art denoising algorithms have made it possible to distinguish individual bacterial taxa at species or strain level, based on single nucleotide polymorphisms (SNP) in the 16S gene (Johnson et al., 2019).

The common conduct of a 16S gene amplicon based taxonomic survey of a microbiome can be stratified in to five main stages; (1) sample collection and storage, (2) DNA extraction, (3) sequencing library preparation and NGS, (4) amplicon sequencing data analysis, (5) data visualization and statistical analysis (Pollock et al., 2018).

Sample collection: Sample collection protocols are largely dependent on the sample type used and the host organism, and can introduce considerable bias to the downstream process. Important aspects to consider here are the sampling site (specific region in gastrointestinal tract of animals or depths in soil), sample collection method (e.g. invasive vs non-invasive), and homogenization of samples (most important in gut and soil studies). Hence, sample collection protocols should always be considered when comparing results of similar studies (Pollock et al., 2018).

Sample storage: The best way for post sample processing is to extract DNA from fresh samples. However, due to practical and logistic issues, this is often impossible and samples will need to be stored for a short or an extended time before DNA can be extracted. Based on current literature the 'gold standard' appears to be to snap freeze samples at sample collection site and store them at -80C^0 until DNA extraction, and most studies report no significant difference between the microbial communities recovered from fresh samples and after extended freezing at -80C^0. Sample storage at 4C^0 however, is not advised and have been demonstrated to have a significant impact on the microbial community recovered (Pollock et al., 2018; Bharti and Grimm, 2019).

DNA Extraction: Existing literature suggests that the DNA extraction protocol used to extract total genomic DNA of samples could introduce considerable bias into the taxonomic classification of a microbial population using 16S gene amplicon sequencing. This is mainly due to microbes that are resistant to lysis, and the presence of inhibitors that may affect the efficiency of the DNA extraction and downstream processes. Hence, the choice of DNA extraction methodology should often be based on the sample type (e.g. fecal, tissue, saliva etc.), inclusion of a mechanical lysis step and the ability to provide a high yield of high-quality DNA. This often requires optimization of current protocols based on DNA extraction kits (Pollock et al., 2018).

Library preparation and sequencing: Sequencing library preparation is also a crucial part of the design of a microbiome study and largely depends on the study objective. If the intent is primarily to do a genus level taxonomic survey of a given microbial population, then generating PCR amplicons of a certain hypervariable region (or a combination) coupled with 2nd generation short read sequencing would suffice. Again, it should be emphasized that there is no current consensus on a particular hypervariable region that can distinguish all bacterial genera with equal efficiency. However, based on existing literature, regions such as V1,V2 and V3,V4 can be identified as commonly used where existing studies have found V3,V4 to over perform V1,V2 (Pollock et al., 2018; Rausch et al., 2019). A study based on geodesic distance has also identified the sub-regions V4-V6 as optimal to identify phylogenetic diversity across bacterial phyla (Yang et al., 2016). Other than the sub-region of choice, factors such as the presence of PCR

inhibitors, the number of PCR cycles, the type of DNA polymerase used in the PCR reaction, and the quantity of input DNA have all been found to significantly impact the structure and composition of the microbial community recovered (Gohl et al., 2016; Pollock et al., 2018).

The phylogenetic analysis of a microbial population at species or strain level, however, would require the entire length of the 16S rRNA gene (~1,500 bps in length) to be included in library preparation, coupled with third generation sequencing such as Oxford Nanopore or PacBio. Techniques such as Circular Consensus Sequencing and complex denoising algorithms that can detect minute PCR and sequencing errors, have made it possible for species and strain level differentiation of 16S data based on Single-nucleotide Polymorphisms (SNPs) (Johnson et al., 2019).

Amplicon sequencing data analysis: 16S amplicon sequencing data are usually generated as 250 bp paired-end reads on Illumina (MiSeq) platforms. The very first step of amplicon sequence data analysis would be to demultiplex the sequencing reads into the respective samples they came from using the barcodes used during sequencing. Subsequently, the paired reads are joined to obtain merged amplicon sequences that are then stripped of their barcodes and linker-primer sequences. Sequence pre-processing usually concludes with a final quality control step where low quality reads are removed based on their quality scores (Pollock et al., 2018; Liu et al., 2020).

Sequences that passes through the above steps will then be used to pick representative sequences for different bacterial species (genera). There are two major approaches used by current analysis pipelines for this purpose; feature selection by clustering into Operational Taxonomic Units (OTUs), and feature selection by de-noising into Amplicon Sequence Variants (ASVs). Algorithms such as VSEARCH and USEARCH can be used to generate OTUs, usually based on >97% sequence similarity. This method however is thought to be less sensitive than the ASV method, currently championed by algorithms such as DADA2 (Callahan et al., 2016), Deblur (Amir et al., 2017) and unoise3 in USEARCH (Edgar, 2016). Post feature selection, the generated OTU/ASV table that contains the frequencies of each representative sequence/OTU (feature) in each of the samples analyzed, is subjected to downstream statistical analysis and data visualizations (Rognes et al., 2016; Liu et al., 2020).

Post feature selection, the representative sequences are aligned against a reference database using a sequence classifier. SILVA, RDP and Greengenes are the most commonly used 16S reference databases for this purpose, and more often than not, SILVA and RDP are preferred over Greengenes due to data quality. The aligned sequences are then assigned with a taxonomic classification. These taxonomic classifications may differ based on the underlying reference database used. The end of the core data analysis process for 16S amplicon data should result in ASV/OTU table and a taxonomy table that contains the taxonomic classification for each ASV/OTU (Pollock et al., 2018).

Data visualization and statistical analysis: Once the core data analysis is completed, a number of statistical analyses and data visualizations can be conducted on the resulting output to generate and extract biologically relevant information. Computational tools such as MEGAN, STAMP, MicrobiomeAnalyst, Phyloseq (a package in R) etc. can be used for this purpose. Some of the common statistical tests include, rarefaction analysis, alpha diversity, beta diversity, Linear Discriminant Analysis Effect Size (LEfSe) etc. (McMurdie and Holmes, 2013; Parks et al., 2014; Callahan et al., 2016; Huson et al., 2016)

1.8.3 Functional profiling of the microbiome through 16S gene amplicon sequencing

16S rRNA gene data are best utilized to obtain information about the taxonomic classification of a microbial community. However, computational tools such as PICRUSt1, PICRUSt2, Tax4Fun and FAPROTAX allow us to gain at least some functional insights into microbial communities by predicting its metagenome using 16S data. The functional inference is done by linking taxonomic classifications with functional data such as KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways, KEGG orthologs and Enzyme Commission numbers (EC numbers) in published databases and literature. Tools such as PICRUSt2 uses advanced phylogenetic tree building algorithms capable of utilizing both OTU and ASV data for functional prediction based on predefined as well as user defined reference databases (Douglas et al., 2020; Liu et al., 2020).

1.9 Whole metagenome sequencing (WMGS) for microbiome studies

Even though 16S amplicon sequencing is considered as a readily available, efficient and cost-effective method of determining the taxonomic composition of a given microbial community, the pitfalls and limitations it inherit, such as PCR amplification bias, variability of results based on the choice of hypervariable region, and inability to accurately classify beyond genus level (unless the entire 16S gene is sequenced) have undermined the accuracy and reproducibility of 16S marker based microbiome studies. Moreover, 16S amplicon sequencing studies are inherently incapable of directly elucidating the functional characteristics of a microbiome, and can only identify taxa for which there are known, amplifiable phylogenetic markers, making it difficult to identify novel organisms. In addition, horizontal gene transfer events that may have occurred in the 16S locus between distant taxa, may give rise to an overestimation of microbial diversity in 16S studies. These well-known shortcomings have made a strong case for the inclusion of Metagenomics or WMGS studies to further analyze microbiomes (Sharpton, 2014; Rausch et al., 2019).

Metagenomics, unlike 16S rRNA sequencing, is not a marker-based approach and attempts to sequence total DNA content in a sample using genome-wide shotgun sequencing. Sample collection and DNA extraction procedure and concerns for WMGS studies are quite similar to 16S studies. Post DNA isolation steps such as library preparation & sequencing, data preprocessing, and sequence data analysis however, is starkly different. That being said, once the primary analysis of sequencing data is

complete, the post-processing analysis that follows has considerable overlap with methods used in amplicon sequence data analyses (Bharti and Grimm, 2019).

Library preparation and sequencing: Illumina sequencing, a popular NGS technology that is commonly used in WMGS studies, offers a variety of library preparation kits that differ based on the method of DNA fragmentation. Commonly used kits such as Nextera DNA Flex, and Nextera XT that utilize bead-based transposomes, and enzymatic lysis for shotgun library preparation, respectively, are known to introduce biases into microbiome studies. Hence, PCR-free library preparation kits such as TruSeq DNA PCR-Free that is based on PCR-free mechanical lysis are generally recommended (Jones et al., 2015; Bharti and Grimm, 2019)

Data pre-processing: Sequence quality checking, sequence quality trimming and decontamination will be done during this step. FastQC, a java-based tool that has the capability of reading FASTQ/SAM/BAM files in order to generate data quality reports on the raw sequencing data can be used to check sequence data quality. This will also be useful to identify which areas need to be trimmed. Once the law quality areas are identified (if at all), a sequence trimming tool like Trimmomatic can be used to remove leading/trailing low quality bases and adaptors. The next crucial step is decontamination via the removal of host DNA sequences using reference based host DNA contaminant removal pipeline such as KneadData (Andrews et al., 2012; Bolger et al., 2014; Liu et al., 2020)

Assembly, binning and taxonomic classification (sequence data analysis): The metagenomic shotgun sequences subjected to preprocessing can be analyzed using either a read-based, assembly-based, or a combinatory approach. Deciding on which approach to use largely depends on the biological question at hand.

If the intent is to determine the species or strain level taxonomic classification of the microbial community, MetaPhlAn can be used on the cleansed metagenomic reads. MetaPhlAn is a metagenomics phylogenetic analysis tool that makes use of a database of clade-specific marker genes obtained from nearly 17,000 reference genomes that span across bacteria, viruses, archaea and eukaryotes to assign taxonomic classifications to metagenome sequence reads. It can provide species level or above taxonomic resolution and a more accurate relative abundance value for each organism, due to its use of single copy genes in the reference database (Segata et al., 2012). A more recent version of the same, MetaPhlAn2 has the ability to do strain level taxonomic identification and strain tracking as well (Truong et al., 2015)**.**

However, this approach cannot classify previously unknown organisms accurately, since it relies on a reference database derived from genomes of known species. Therefore, as a further analysis step, sequence assembly and binning should be done to arrive at a metagenome assembled genome or MAG. The unique nature of whole metagenome shotgun sequences presents a unique set of challenges for assemblers trying to assemble them into scaffolds and contigs. Therefore, specialized de-novo assemblers are required to overcome them. IDBA-UD, MEGAHIT and MetaSPAdes are such metagenome

specific assemblers suited for this task. However, no matter which assembler is used, the output would be a set of contigs. Hence, specialized software is required to bin these contigs into genomes (Quince et al., 2017; Olson et al., 2018).

Binning methods can be grouped into two broad categories; supervised and unsupervised. Supervised methods are designed to bin contigs based on reference databases and unsupervised methods are designed to find natural groups based on clustering of similar contigs. Methods that use supervised binning is not often considered as effective when analyzing environmental samples that can be home to diverse microbiota (Quince et al., 2017; Woloszynek et al., 2018). Tools such as MaxBin and MetaCluster5.0 can be used for binning of assembled contigs(Sedlar et al., 2017). Once binning is completed, tools such as CheckM and Taxator-tk can be used for assigning taxonomic classifications to the binned sequences (Dröge et al., 2015; Parks et al., 2015).

Functional profiling (data analysis): Identifying genes and metabolic pathways using whole metagenome sequencing data accurately reveals the functional profile and metabolic potential of a given microbial community. This can be done either using the MAG assemblies described above, or directly using the pre-processed shotgun sequence reads (Quince et al., 2017; Liu et al., 2020).

HUMAnN2 is popular tool used for this purpose that uses pre-processed shotgun sequences as the input. HUMAnN2 relies on a tiered search strategy that consists of identifying known players in the community via marker genes and constructing a

pangenome, mapping shotgun sequence reads to each pangenome, and translating unmapped sequence reads for protein database similarity searches. This tiered approach gives it the ability to ultimately analyze all sequences with known or unknown taxonomy. At the end of this process HUMAnN2 will be able to provide a comprehensive visual analysis on abundance of gene families, abundance of metabolic pathways and pathway coverage (Franzosa et al., 2018).

For an assembly-based approach, a software pipeline such as Prokka can be used to annotate the MAGs. Prokka is a Unix based software that can be used to annotate shotgun sequences that are assembled into scaffolds. It uses external computational tools such as Prodigal, RNAmmer, Aragorn, SignalP and Infernal to annotate genomic features and their coordinates on pre-assembled contigs. It utilizes parallel programming techniques to exploit multi-core processor for efficient analysis of bacterial genomic data (Seemann, 2014).

Post-processing analyses: Similar to 16S studies, once the core sequence data analysis steps are complete, the WMGS studies will also end-up with matrices of samples against taxa, genes, pathways etc. Hence, most of statistical analysis and data visualization methods and tools used for amplicon sequencing (16S) studies can be used for metagenome studies as well. Post-processing analyses can be divided into two categories; supervised and unsupervised (Quince et al., 2017).

Supervised methods would include the traditional multivariate analysis techniques (e.g. ANOVA), and machine learning algorithms that classify samples into groups based on a training dataset. Unsupervised methods would include hierarchical clustering of samples and heatmaps, sample clustering based on principle component analysis (PCA) and principle coordinate analysis (PCoA), and pattern detection via correlation network analysis. Indices that measure between sample differences (beta diversity), and species richness and evenness (alpha diversity), can be used to assess the overall composition of a feature table along with abundance graphs and box plots for data visualization. Computational tools such as MicrobiomeAnalyst, MEGAN and R packages such as DESeq, metagenomeSeq, Cytoscape and vegan are often utilized to conduct these statistical analyses and data visualizations (Huson et al., 2016; Dhariwal et al., 2017; Quince et al., 2017; Liu et al., 2020).

The read based approach was utilized in metagenome sequence data analysis for the current study. Hence, MetaPhlAn 3.0 (Truong et al., 2015) was used for taxonomic profiling of metagenome sequencing data and HUMAnN 3.0 (Franzosa et al., 2018) was used for functional profiling of rumen and fecal samples.

1.10 References

Ames, N. J., A. Ranucci, B. Moriyama, and G. R. Wallen. 2017. The Human Microbiome and Understanding the 16S rRNA Gene in Translational Nursing Science. Nurs. Res. 66:184–197. doi:10.1097/NNR.0000000000000212.

Amir, A., M. Daniel, J. Navas-Molina, E. Kopylova, J. Morton, Z. Z. Xu, K. Eric, L. Thompson, E. Hyde, A. Gonzalez, and R. Knight. 2017. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. J. A. Gilbert, editor. mSystems. 2:1–7. doi:10.1128/mSystems.00191-16.

Anderson, C. L., M. B. Sullivan, and S. C. Fernando. 2017. Dietary energy drives the dynamic response of bovine rumen viral communities. Microbiome. 5:155. doi:10.1186/s40168-017-0374-3.

Andrews, S., F. Krueger, A. Segonds-Pichon, L. Biggins, C. Krueger, and S. Wingett. 2012. FastQC: a quality control tool for high throughput sequence data.

An Overview of the U.S. Beef Industry. Available from: https://www.pearsonhighered.com/assets/samplechapter/0/1/3/4/0134602692.pdf

Baek, Y. C., H. Choi, J.-Y. Jeong, S. D. Lee, M. J. Kim, S. Lee, S.-Y. Ji, and M. Kim. 2020. The impact of short-term acute heat stress on the rumen microbiome of Hanwoo steers. J. Anim. Sci. Technol. 62:208–217. doi:10.5187/jast.2020.62.2.208.

Berg, G., D. Rybakova, D. Fischer, T. Cernava, M. C. C. Vergès, T. Charles, X. Chen, L. Cocolin, K. Eversole, G. H. Corral, M. Kazou, L. Kinkel, L. Lange, N. Lima, A. Loy, J. A. Macklin, E. Maguin, T. Mauchline, R. McClure, B. Mitter, M. Ryan, I. Sarand, H. Smidt, B. Schelkle, H. Roume, G. S. Kiran, J. Selvin, R. S. C. de Souza, L. Van Overbeek, B. K. Singh, M. Wagner, A. Walsh, A. Sessitsch, and M. Schloter. 2020. Microbiome definition re-visited: old concepts and new challenges. Microbiome. 8:1–22. doi:10.1186/s40168-020-00875-0.

Bernabucci, U. 2019. Climate change: impact on livestock and how can we adapt. Anim. Front. 9:3–5. doi:10.1093/af/vfy039.

Bharti, R., and D. G. Grimm. 2019. Current challenges and best-practice protocols for microbiome analysis. Brief. Bioinform. 22:178–193. doi:10.1093/bib/bbz155.

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

Bowling, G. A. 1942. The Introduction of Cattle into Colonial North America. J. Dairy Sci. 25:129–154. doi:10.3168/jds.S0022-0302(42)95275-5.

Brulc, J. M., D. A. Antonopoulos, M. E. Berg Miller, M. K. Wilson, A. C. Yannarell, E. A. Dinsdale, R. E. Edwards, E. D. Frank, J. B. Emerson, P. Wacklin, P. M. Coutinho, B.Henrissat, K. E. Nelson, and B. A. White. 2009. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. Proc. Natl. Acad. Sci. U. S. A. 106:1948–1953. doi:10.1073/pnas.0806191105.

Callahan, Benjamin J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583. doi:10.1038/nmeth.3869.

Chakravorty, S., D. Helb, M. Burday, N. Connell, and D. Alland. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J. Microbiol. Methods. 69:330–9. doi:10.1016/j.mimet.2007.02.005.

Chaucheyras-Durand, F., and F. Ossa. 2014. The rumen microbiome: Composition, abundance, diversity, and new investigative tools. Prof. Anim. Sci. 30:1–12. doi:10.15232/S1080-7446(15)30076-0.

Clauss, M., I. D. Hume, and J. Hummel. 2010. Evolutionary adaptations of ruminants and their potential relevance for modern production systems. Animal. 4:979–992. doi:10.1017/S1751731110000388.

Countryman, A. M., P. L. Paarlberg, and J. G. Lee. 2016. Dynamic effects of drought on the U.S. beef supply chain. Agric. Resour. Econ. Rev. 45:459–484. doi:10.1017/age.2016.4.

Delgado, B., A. Bach, I. Guasch, C. González, G. Elcoso, J. E. Pryce, and O. Gonzalez-Recio. 2019. Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake levels in cattle. Sci. Rep. 9:1–13. doi:10.1038/s41598-018-36673-w.

DePeters, E. J., and L. W. George. 2014. Rumen transfaunation. Immunol. Lett. 162:69–76. doi:10.1016/j.imlet.2014.05.009. 9

Dhariwal, A., J. Chong, S. Habib, I. L. King, L. B. Agellon, and J. Xia. 2017. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 45:W180–W188. doi:10.1093/nar/gkx295.

Douglas, G. M., V. J. Maffei, J. R. Zaneveld, S. N. Yurgel, J. R. Brown, C. M. Taylor, C. Huttenhower, and M. G. I. Langille. 2020. PICRUSt2 for prediction of metagenome functions. Nat. Biotechnol. 38:685–688. doi:10.1038/s41587-020-0548-6.

Dröge, J., I. Gregor, and A. C. McHardy. 2015. Taxator-tk: Precise taxonomic assignment of metagenomes by fast approximation of evolutionary neighborhoods. Bioinformatics. 31:817–824. doi:10.1093/bioinformatics/btu745.

Drouillard, J. S. 2018. Current situation and future trends for beef production in the United States of America - A review. Asian-Australasian J. Anim. Sci. 31:1007–1016. doi:10.5713/ajas.18.0428.

Dyer, T., J. Rossy, and M. Pence. 2017. Water Requirements and Quality Issues for Cattle. Available from: https://secure.caes.uga.edu/extension/publications/files/pdf/SB 56_5.PDF

Easterling, D. R., K. E. Kunkel, J. R. Arnold, T. R. Knutson, A. N. LeGrande, L. R. Leung, R. S. Vose, D. E. Waliser, and M. Wehner. 2017. Precipitation change in the United States. In: Climate Science Special Report: Fourth National Climate Assessment. Fourth Natl. Clim. Assessment, Vol. I. I:207–230. doi:10.7930/J0H993CC.U.S.

Edgar, R. 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv. 081257. doi:10.1101/081257.

Fan, P., B. Bian, L. Teng, C. D. Nelson, J. Driver, M. A. Elzo, and K. C. Jeong. 2020. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 14:302–317. doi:10.1038/s41396-019-0529-2.

FAO. 2013. Greenhouse gas emissions from ruminant supply chains. Available from: http://www.fao.org/docrep/018/i3461e/i3461e00.htm

Ferreira, L. M. M., G. Hervás, A. Belenguer, R. Celaya, M. A. M. Rodrigues, U. García, P. Frutos, and K. Osoro. 2017. Comparison of feed intake, digestion and rumen function among domestic ruminant species grazing in upland vegetation communities. J. Anim. Physiol. Anim. Nutr. (Berl). 101:846–856. doi:10.1111/jpn.12474.

Fouts, D. E., S. Szpakowski, J. Purushe, M. Torralba, R. C. Waterman, M. D. MacNeil, L. J. Alexander, and K. E. Nelson. 2012. Next Generation Sequencing to Define Prokaryotic and Fungal Diversity in the Bovine Rumen. PLoS One. 7. doi:10.1371/journal.pone.0048289.

Frandson, R. D., W. L. Wilke, and A. D. Fails. 2019. Anatomy and Physiology of Farm Animals seventh. 7th Ed. Wiley-Blackwell, Hoboken, NJ. p. 346–350.

Franzosa, E. A., L. J. McIver, G. Rahnavard, L. R. Thompson, M. Schirmer, G. Weingart, K. S. Lipson, R. Knight, J. G. Caporaso, N. Segata, and C. Huttenhower. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. Nat. Methods. 15:962–968. doi:10.1038/s41592-018-0176-y. 176-y

Gerber, P. J., A. Mottet, C. I. Opio, A. Falcucci, and F. Teillard. 2015. Environmental impacts of beef production: Review of challenges and perspectives for durability. Meat Sci. 109:2–12. doi:10.1016/j.meatsci.2015.05.013.

Gilbert, R. A., E. M. Townsend, K. S. Crew, T. C. A. Hitch, J. C. A. Friedersdorff, C. J. Creevey, P. B. Pope, D. Ouwerkerk, and E. Jameson. 2020. Rumen Virus Populations: Technological Advances Enhancing Current Understanding. Front. Microbiol. 11:450. doi:10.3389/fmicb.2020.00450.

Gohl, D. M., P. Vangay, J. Garbe, A. MacLean, A. Hauge, A. Becker, T. J. Gould, J. B. Clayton, T. J. Johnson, R. Hunter, D. Knights, and K. B. Beckman. 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat. Biotechnol. 34:942–949. doi:10.1038/nbt.3601.

Guo, W., W. Wang, S. Bi, R. Long, F. Ullah, M. Shafiq, M. Zhou, and Y. Zhang. 2020.

Characterization of anaerobic rumen fungal community composition in Yak, Tibetan sheep and small tail han sheep grazing on the Qinghai-Tibetan Plateau. Animals. 10. doi:10.3390/ani10010144.

Gupta, S., M. S. Mortensen, S. Schjørring, U. Trivedi, G. Vestergaard, J. Stokholm, H. Bisgaard, K. A. Krogfelt, and S. J. Sørensen. 2019. Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. Commun. Biol. 2:1–7. doi:10.1038/s42003-019-0540-1.

Hackmann, T. J., and J. L. Firkins. 2015. Maximizing efficiency of rumen microbial protein production. Front. Microbiol. 6:1–16. doi:10.3389/fmicb.2015.00465.

Hackmann, T. J., and J. N. Spain. 2010. Invited review: Ruminant ecology and evolution: Perspectives useful to ruminant livestock research and production. J. Dairy Sci. 93:1320–1334. doi:10.3168/jds.2009-2071.

Hanafy, R. A., B. Johnson, N. H. Youssef, and M. S. Elshahed. 2020. Assessing anaerobic gut fungal diversity in herbivores using D1/D2 large ribosomal subunit sequencing and multi-year isolation. Environ. Microbiol. 22:3883–3908. doi:10.1111/1462-2920.15164.

Heather, J. M., and B. Chain. 2016. The sequence of sequencers: The history of sequencing DNA. Genomics. 107:1–8. doi:10.1016/j.ygeno.2015.11.003.

Heikema, A. P., D. Horst-Kreft, S. A. Boers, R. Jansen, S. D. Hiltemann, W. de Koning, R. Kraaij, M. A. J. de Ridder, C. B. van Houten, L. J. Bont, A. P. Stubbs, and J. P. Hays. 2020. Comparison of illumina versus nanopore 16s rRNA gene sequencing of the human nasal microbiota. Genes (Basel). 11:1–17. doi:10.3390/genes11091105.

Henderson, G., F. Cox, S. Ganesh, A. Jonker, W. Young, P. H. Janssen, L. Abecia, E. Angarita, P. Aravena, G. N. Arenas, C. Ariza, G. T. Attwood, J. M. Avila, J. Avila-Stagno, A. Bannink, R. Barahona, M. Batistotti, M. F. Bertelsen, A. Brown-Kav, A. M. Carvajal, L. Cersosimo, A. V. Chaves, J. Church, N. Clipson, M. A. Cobos-Peralta, A. L. Cookson, S. Cravero, O. C. Carballo, K. Crosley, G. Cruz, M. C. Cucchi, R. De La Barra, A. B. De Menezes, E. Detmann, K. Dieho, J. Dijkstra, W. L. S. Dos Reis, M. E. R. Dugan, S. H. Ebrahimi, E. Eythórsdóttir, F. N. Fon, M. Fraga, F. Franco, C. Friedeman, N. Fukuma, D. Gagić, I. Gangnat, D. J. Grilli, L. L. Guan, V. H. Miri, E. Hernandez-Sanabria, A. X. I. Gomez, O. A. Isah, S. Ishaq, E. Jami, J. Jelincic, J. Kantanen, W. J. Kelly, S. H. Kim, A. Klieve, Y. Kobayashi, S. Koike, J. Kopecny, T. N. Kristensen, S. J. Krizsan, H. LaChance, M. Lachman, W. R. Lamberson, S. Lambie, J. Lassen, S. C. Leahy, S. S. Lee, F. Leiber, E. Lewis, B. Lin, R. Lira, P. Lund, E. Macipe, L. L. Mamuad, H. C. Mantovani, G. A. Marcoppido, C. Márquez, C. Martin, G. Martinez, M. E. Martinez, O. L. Mayorga, T. A. McAllister, C. McSweeney, L. Mestre, E. Minnee, M. Mitsumori, I. Mizrahi, I. Molina, A. Muenger, C. Munoz, B. Murovec, J. Newbold, V. Nsereko, et al. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5. doi:10.1038/srep14567.

Huson, D. H., S. Beier, I. Flade, A. Górska, M. El-Hadidi, S. Mitra, H. J. Ruscheweyh, and R. Tappu. 2016. MEGAN Community Edition - Interactive Exploration and Analysis

of Large-Scale Microbiome Sequencing Data. PLoS Comput. Biol. 12:1–12. doi:10.1371/journal.pcbi.1004957.

Huws, S. A., C. J. Creevey, L. B. Oyama, I. Mizrahi, S. E. Denman, M. Popova, R. Muñoz-Tamayo, E. Forano, S. M. Waters, M. Hess, I. Tapio, H. Smidt, S. J. Krizsan, D. R. Yáñez-Ruiz, A. Belanche, L. Guan, R. J. Gruninger, T. A. McAllister, C. J. Newbold, R. Roehe, R. J. Dewhurst, T. J. Snelling, M. Watson, G. Suen, E. H. Hart, A. H. Kingston-Smith, N. D. Scollan, R. M. Do Prado, E. J. Pilau, H. C. Mantovani, G. T. Attwood, J. E. Edwards, N. R. McEwan, S. Morrisson, O. L. Mayorga, C. Elliott, and D. P. Morgavi. 2018. Addressing global ruminant agricultural challenges through understanding the rumen microbiome: Past, present, and future. Front. Microbiol. 9:1–33. doi:10.3389/fmicb.2018.02161.

Ishler, V., J. Heinrichs, and G. Varga. 1996. From Feed to Milk: Understanding Rumen Function. Ext. Circ. 422:1–32.

Johnson, J. S., D. J. Spakowicz, B. Y. Hong, L. M. Petersen, P. Demkowicz, L. Chen, S. R. Leopold, B. M. Hanson, H. O. Agresta, M. Gerstein, E. Sodergren, and G. M. Weinstock. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat. Commun. 10:1–11. doi:10.1038/s41467-019-13036-1.

Jones, M. B., S. K. Highlander, E. L. Anderson, W. Li, M. Dayrit, N. Klitgord, M. M. Fabani, V. Seguritan, J. Green, D. T. Pride, S. Yooseph, W. Biggs, K. E. Nelson, and J. Craig Venter. 2015. Library preparation methodology can influence genomic and functional predictions in human microbiome research. Proc. Natl. Acad. Sci. U. S. A. 112:14024–14029. doi:10.1073/pnas.1519288112.

Kellogg, E. A. 2001. Update on Evolution Evolutionary History of the Grasses. Plant Physiol. 125:1198–1205. doi:10.1104/pp.125.3.1198.

Kim, M., and J. Chun. 2014. 16S rRNA gene-based identification of bacteria and archaea using the EzTaxon server. In: Methods in Microbiology. Vol. 41. 1st ed. Elsevier Ltd. p. 61–74.

Krause, D. O., T. G. Nagaraja, A. D. G. Wright, and T. R. Callaway. 2013. Board-invited review: Rumen microbiology: Leading the way in microbial ecology. J. Anim. Sci. 91:331–341. doi:10.2527/jas.2012-5567.

Li, F., C. Li, Y. Chen, J. Liu, C. Zhang, B. Irving, C. Fitzsimmons, G. Plastow, and L. L. Guan. 2019. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome. 7:1–17. doi:10.1186/s40168-019-0699-1.

Liggenstoffer, A. S., N. H. Youssef, M. B. Couger, and M. S. Elshahed. 2010. Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. ISME J. 4:1225–1235. doi:10.1038/ismej.2010.49.

Liu, Y. X., Y. Qin, T. Chen, M. Lu, X. Qian, X. Guo, and Y. Bai. 2020. A practical guide to amplicon and metagenomic analysis of microbiome data. Protein Cell. doi:10.1007/s13238-020-00724-8.

Loor, J. J., A. A. Elolimy, and J. C. McCann. 2016. Dietary impacts on rumen microbiota in beef and dairy production. Anim. Front. 6:22–29. doi:10.2527/af.2016-0030.

Lu, H., F. Giordano, and Z. Ning. 2016. Oxford Nanopore MinION Sequencing and Genome Assembly. Genomics, Proteomics Bioinformatics. 14:265–279. doi:10.1016/j.gpb.2016.05.004.

Marchesi, J. R., and J. Ravel. 2015. The vocabulary of microbiome research: a proposal. Microbiome. 3:1–3. doi:10.1186/s40168-015-0094-5.

Matthews, C., F. Crispie, E. Lewis, M. Reid, P. W. O'Toole, and P. D. Cotter. 2019. The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. Gut Microbes. 10:115–132. doi:10.1080/19490976.2018.1505176.

Mccann, J. C., T. A. Wickersham, and J. J. Loor. 2014. High-throughput Methods Redefine the Rumen Microbiome and Its Relationship with Nutrition and Metabolism. Bioinform. Biol. Insights. 8:BBI.S15389. doi:10.4137/bbi.s15389.

McMurdie, P. J., and S. Holmes. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One. 8. doi:10.1371/journal.pone.0061217.

McSweeney, C. S., and R. Mackie. 2012. Mirco-organisms and ruminant digestion: state of knowledge, trends and future prospects. Backgr. study Pap. 62.

Moss, E. L., D. G. Maghini, and A. S. Bhatt. 2020. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. Nat. Biotechnol. 38:701–707. doi:10.1038/s41587-020-0422-6.

Namonyo, S., M. Wagacha, S. Maina, L. Wambua, and M. Agaba. 2018. A metagenomic study of the rumen virome in domestic caprids. Arch. Virol. 163:3415–3419. doi:10.1007/s00705-018-4022-4.

Nardone, A., B. Ronchi, N. Lacetera, M. S. Ranieri, and U. Bernabucci. 2010a. Effects of climate changes on animal production and sustainability of livestock systems. Livest. Sci. 130:57–69. doi:10.1016/j.livsci.2010.02.011.

Olson, N. D., T. J. Treangen, C. M. Hill, V. Cepeda-Espinoza, J. Ghurye, S. Koren, and M. Pop. 2018. Metagenomic assembly through the lens of validation: Recent advances in assessing and improving the quality of genomes assembled from metagenomes. Brief. Bioinform. 20:1140–1150. doi:10.1093/bib/bbx098.

Overpeck, J. T., and B. Udall. 2020. Climate change and the aridification of North America. Proc. Natl. Acad. Sci. 117:11856–11858. doi:10.1073/pnas.2006323117.

Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson. 2015. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 25:1043–1055. doi:10.1101/gr.186072.114.

Parks, D. H., G. W. Tyson, P. Hugenholtz, and R. G. Beiko. 2014. STAMP: Statistical analysis of taxonomic and functional profiles. Bioinformatics. 30:3123–3124.

doi:10.1093/bioinformatics/btu494.

Derell, P. S. Economic Damages to the U.S. Beef Cattle Industry Due to COVID-19. 2020.

Pitta, D. W., W. E. Pinchak, S. E. Dowd, J. Osterstock, V. Gontcharova, E. Youn, K. Dorton, I. Yoon, B. R. Min, J. D. Fulford, T. A. Wickersham, and D. P. Malinowski. 2010. Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microb. Ecol. 59:511–522. doi:10.1007/s00248-009-9609-6.

Pollock, J., L. Glendinning, T. Wisedchanwet, and M. Watson. 2018. The Madness of Microbiome : Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. Appl. Environ. Microbiol. 1–12.

Quince, C., A. W. Walker, J. T. Simpson, N. J. Loman, and N. Segata. 2017. Shotgun metagenomics, from sampling to analysis. Nat. Biotechnol. 35:833–844. doi:10.1038/nbt.3935.

Rasby, R. J., and T. M. Walz. 2011. Water Requirements for Beef Cattle.

Rausch, P., M. Rühlemann, B. M. Hermes, S. Doms, T. Dagan, K. Dierking, H. Domin, S. Fraune, J. Von Frieling, U. Hentschel, F. A. Heinsen, M. Höppner, M. T. Jahn, C. Jaspers, K. A. B. Kissoyan, D. Langfeldt, A. Rehman, T. B. H. Reusch, T. Roeder, R. A. Schmitz, H. Schulenburg, R. Soluch, F. Sommer, E. Stukenbrock, N. Weiland-Bräuer, P. Rosenstiel, A. Franke, T. Bosch, and J. F. Baines. 2019b. Comparative analysis of amplicon and metagenomic sequencing methods reveals key features in the evolution of animal metaorganisms. Microbiome. 7:1–19. doi:10.1186/s40168-019-0743-1.

Robb, A., and Z. Zhao. 2015. An Overview of the U.S. Beef Industry. 15:9–21.

Roehe, R., R. J. Dewhurst, C. A. Duthie, J. A. Rooke, N. McKain, D. W. Ross, J. J. Hyslop, A. Waterhouse, T. C. Freeman, M. Watson, and R. J. Wallace. 2016. Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. PLoS Genet. 12:1–20. doi:10.1371/journal.pgen.1005846.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ. 2016:1–22. doi:10.7717/peerj.2584.

Romero-Pérez, G. A., K. H. Ominski, T. A. McAllister, and D. O. Krause. 2011. Effect of environmental factors and influence of rumen and hindgut biogeography on bacterial communities in steers. Appl. Environ. Microbiol. 77:258–268. doi:10.1128/AEM.01289-09.

Sedlar, K., K. Kupkova, and I. Provaznik. 2017. Bioinformatics strategies for taxonomy independent binning and visualization of sequences in shotgun metagenomics. Comput. Struct. Biotechnol. J. 15:48–55. doi:10.1016/j.csbj.2016.11.005.

Seemann, T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 30:2068–2069. doi:10.1093/bioinformatics/btu153.

Segata, N., L. Waldron, A. Ballarini, V. Narasimhan, O. Jousson, and C. Huttenhower. 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat. Methods. 9:811–814. doi:10.1038/nmeth.2066.

Sequencing 101: The Evolution of DNA Sequencing Tools - PacBio. Available from: https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/

Seshadri, R., S. C. Leahy, G. T. Attwood, K. H. Teh, S. C. Lambie, A. L. Cookson, E. A. Eloe-Fadrosh, G. A. Pavlopoulos, M. Hadjithomas, N. J. Varghese, D. Paez-Espino, R. Perry, G. Henderson, C. J. Creevey, N. Terrapon, P. Lapebie, E. Drula, V. Lombard, E. Rubin, N. C. Kyrpides, B. Henrissat, T. Woyke, N. N. Ivanova, W. J. Kelly, N. Palevic, P. H. Janssen, R. S. Ronimus, S. Noel, P. Soni, K. Reilly, T. Atherly, C. Ziemer, A. D. Wright, S. Ishaq, M. Cotta, S. Thompson, K. Crosley, N. McKain, J. J. Wallace, H. J. Flint, J. C. Martin, R. J. Forster, R. J. Gruninger, T. McAllister, R. Gilbert, D. J. Ouwerkerk, A. J. Klieve, R. Al Jassim, S. Denman, C. McSweeney, C. Rosewarne, S. Koike, Y. Kobayashi, M. Mitsumori, T. Shinkai, S. Cravero, and M. Cerón Cucchi. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. Nat. Biotechnol. 36:359–367. doi:10.1038/nbt.4110.

Sharpton, T. J. 2014. An introduction to the analysis of shotgun metagenomic data. Front. Plant Sci. 5:1–14. doi:10.3389/fpls.2014.00209.

Singh, R. R. 2017. Next generation sequencing technologies. Compr. Med. Chem. III. 2–8:354–361. doi:10.1016/B978-0-12-409547-2.12327-3.

Skillman, L. C., A. F. Toovey, A. J. Williams, and A. D. G. Wright. 2006. Development and validation of a real-time PCR method to quantify rumen protozoa and examination of variability between Entodinium populations in sheep offered a hay-based diet. Appl. Environ. Microbiol. 72:200–206. doi:10.1128/AEM.72.1.200-206.2006.

Söllinger, A., A. T. Tveit, M. Poulsen, S. J. Noel, M. Bengtsson, J. Bernhardt, A. L. Frydendahl Hellwing, P. Lund, K. Riedel, C. Schleper, O. Højberg, and T. Urich. 2018. Holistic Assessment of Rumen Microbiome Dynamics through Quantitative Metatranscriptomics Reveals Multifunctional Redundancy during Key Steps of Anaerobic Feed Degradation. mSystems. 3:1–19. doi:10.1128/msystems.00038-18.

Stoddard, S. F., B. J. Smith, R. Hein, B. R. K. Roller, and T. M. Schmidt. 2015. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic Acids Res. 43:593–598. doi:10.1093/nar/gku1201.

Summer, A., I. Lora, P. Formaggioni, and F. Gottardo. 2019. Impact of heat stress on milk and meat production. Anim. Front. 9:39–46. doi:10.1093/af/vfy026.

Sylvester, J. T., S. K. R. Karnati, Z. Yu, M. Morrison, and J. L. Firkins. 2004. Development of an assay to quantify rumen ciliate protozoal biomass in cows using real-time PCR. J. Nutr. 134:3378–3384. doi:10.1093/jn/134.12.3378.

Tajima, K., I. Nonaka, K. Higuchi, N. Takusari, M. Kurihara, A. Takenaka, M. Mitsumori, H. Kajikawa, and R. I. Aminov. 2007. Influence of high temperature and humidity on rumen bacterial diversity in Holstein heifers. Anaerobe. 13:57–64.

doi:10.1016/j.anaerobe.2006.12.001.

Truong, D. T., E. A. Franzosa, T. L. Tickle, M. Scholz, G. Weingart, E. Pasolli, A. Tett, C. Huttenhower, and N. Segata. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods. 12:902–903. doi:10.1038/nmeth.3589.

USDA ERS - Sector at a Glance. 2018. USDA ERS. Available from: https://www.ers.usda.gov/topics/animal-products/cattle-beef/sector-at-a-glance/

USGCRP. 2017. Executive summary. (D. J. Wuebbles, D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart, and T. K. Maycock, editors.). U.S. Global Change Research Program. Available from: https://science2017.globalchange.gov/chapter/executive-summary/

Vose, R. S., D. R. Easterling, K. E. Kunkel, A. N. LeGrande, and M. F. Wehner. 2017. Temperature changes in the United States. Clim. Sci. Spec. Rep. Fourth Natl. Clim. Assess. I:185–206. doi:10.7930/J0N29V45.U.S.

Weimer, P. J., J. B. Russell, and R. E. Muck. 2009. Lessons from the cow: What the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass. Bioresour. Technol. 100:5323–5331. doi:10.1016/j.biortech.2009.04.075.5

Wickramasinghe, H. K. J. P., J. M. Anast, S. Schmitz-Esser, N. V. L. Serão, and J. A. D. R. N. Appuhamy. 2020. Beginning to offer drinking water at birth increases the species richness and the abundance of Faecalibacterium and Bifidobacterium in the gut of preweaned dairy calves. J. Dairy Sci. 103:4262–4274. doi:10.3168/jds.2019-17258.

Wickramasinghe, H. K. J. P., A. J. Kramer, and J. A. D. R. N. Appuhamy. 2019. Drinking water intake of newborn dairy calves and its effects on feed intake, growth performance, health status, and nutrient digestibility. J. Dairy Sci. 102:377–387. doi:10.3168/jds.2018-15579.

Wilson, L. L., K. G. Macdonald, and H. H. Mayo. 1965. Development of Beef Cattle Industry. Available from: https://docs.lib.purdue.edu/anrhist/3

Woloszynek, S., Z. Zhao, G. Ditzler, J. R. Price, R. Reichenberger, Y. Lan, J. Chen, J. Earl, K. Langroodi, G. Ehrlich, and G. Rosen. 2018. Theoretical and Applied Aspects of Systems Biology. 7th Edition. Springer International Publishing, Cham, Switzerland.

Yang, B., Y. Wang, and P. Y. Qian. 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. BMC Bioinformatics. 17:1–8. doi:10.1186/s12859-016-0992-y.

Yarza, P., P. Yilmaz, E. Pruesse, F. O. Glöckner, W. Ludwig, K. H. Schleifer, W. B. Whitman, J. Euzéby, R. Amann, and R. Rosselló-Móra. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat. Rev. Microbiol. 12:635–645. doi:10.1038/nrmicro3330.

Zhao, S., L. Min, N. Zheng, and J. Wang. 2019. Effect of heat stress on bacterial composition and metabolism in the rumen of lactating dairy cows. Animals. 9. doi:10.3390/ani9110925.

CHAPTER II

POPULATION DYNAMICS OF THE BOVINE RUMEN MICROBIOME UNDER WATER RESTRICTED CONDITIONS: A TAXONOMIC SURVEY OF 16S V4 AMPLICON SEQUECING DATA

2.1. Introduction

Even though overlooked in many of the nutrient requirement models, water is one of the most important nutrients required for the health and wellbeing of beef cattle. Water on average, accounts for 50% to 80% of an animal's body weight (BW), and plays an imperative role in feed digestion and maintaining other critical physiological functions (Dyer et al. 2017). Hence, if an animal is deprived of sufficient amounts of water, its feed intake may also drastically decrease, and even bodily functions may fail due to dehydration (Wickramasinghe et al. 2019).

The minimum amount of water required by cattle to properly maintain its physiological functions is determined by growth, water lost due to urinary excretions, sweat, feces and evaporation (through lungs or skin). Furthermore, the daily water requirement of an animal can be influenced by body weight, ambient temperature, stage of production, diet and the type of cattle (Dyer et al. 2017; Rasby and Walz 2011).

Decreased water intake is known to amplify the effects of heat stress in cattle, thus lowering digesta flow rate in the rumen, decreasing saliva excretion, altering feed

efficiency & nutrient digestibility and affecting the composition of blood hormones (Benatallah et al. 2019). Water restriction is known to have a direct effect on rumen function where increased osmolality in rumen fluid caused by water restriction can give rise to dehydration-induced hypophagia (Burgos et al. 2000). Furthermore, a combination of reduced feed intake due to water restriction, reduced saliva production and a decrease in HCO_3^- content in saliva in a heat stressed animal, can make it more susceptible to rumen acidosis (Nardone et al., 2010). Although no direct effects of water restriction on rumen motility is known, an increase in the time to first rumination have been observed (Carter and Grovum, 1990).

While the impact of water restriction on rumen physiology is well known, its impact on the rumen microbiome has not been extensively studied yet, even though rumen microbial studies have seen drastic improvements due to the use of Next Generation Sequencing (NGS) technologies. Existing studies have largely focused on the effect of host genetics (Roehe et al., 2016; Li et al., 2019; Fan et al., 2020), age (Liu et al., 2017), sex (Li et al., 2019), feed intake (Delgado *et al.*, 2019), diet (Pitta *et al.*, 2010), and heat stress (Tajima et al., 2007; Romero-Pérez et al., 2011; Zhao et al., 2019; Baek et al., 2020) on the rumen microbial population, with meager emphasis on how it may be affected by water scarcity or water restriction.

Hence, this study was conducted to elucidate the microbial dynamics of the bovine rumen under water stress, primarily based on 16S rRNA gene amplicon sequencing of genomic

DNA extracted from rumenocentesis samples obtained from feedlot beef cattle under 50% water restriction.

2.2. Materials and Methods

2.2.1.Animals

A total of 858 cross-bred steers belonging to 7 feeding groups were used in a randomized block design for this study. Group 01 was maintained from May 2014 to October 2014 (n=117), group 02 from November 2014 to May 2015 (n=115), group 03 from May 2015 to September 2015 (n=118), group 05 from May 2016 to November 2016 (n=105), group 06 from December 2016 to June 2017 (n=123), group 07 from August 2017 to January 2018 (n=142) and group 08 from February 2018 to July 2018 (n=138). All animals were housed in the Willard Sparks Beef Research Center (WSBRC) in Stillwater, Oklahoma.

Before pen assignments (on d 21), all animals were vaccinated for bacterial (Vision 7®, Merck Animal Health, Madison, NJ) and viral (Titanium 5+PH-M, Elanco, Greenfield, IN) diseases. Administering fenbendazole oral suspension (Safe-Guard by Merck Animal Health), and metaphylaxis treatment (Excede® by Zoetis) was also conducted to deworm and prevent the spread of infectious diseases. The steers were also implanted with an estradiol implant (Compudose® by Elanco) upon first arrival (Bruno, 2019). The animals were fed with a diet comprising of ~51% wet corn gluten feed, 15% cracked corn, ~28% prairie hay, and ~5% feed additives (supplements). Animal breeds were largely of British influence and were proactively monitored to exclude *Bos indicus* descendants,

since they are known to be capable of intentionally reducing water consumption as an adaptation to increased ambient temperature (Ahlberg, 2018).

2.2.2. Experimental design and sample collection

Within each group, the animals were randomly allocated to one of 4 pens (approx. 30 animals each) in a completely randomized design. All animals were given a 21-day acclimatization period upon arrival at the WSBRC. A 70-day feed and water intake (WI) trial followed after the acclimatization period to establish baseline WI. All animals had access to *ad libitum* feed and water during the acclimatization period and WI trial.

Post WI trial, 50% water restriction was achieved via gradual reduction of water across a 35-day step-down period, giving animals 7 days to get acclimatized to a 10% decrease in available water. Animals were maintained at 50% restriction for another 35 days to complete the water restriction trial, before being brought up to *ad libitum* WI levels within 6 days of ending the water restriction. Insentec® (Hokofarm Group, Netherlands) Roughage Intake Control (RIC) system was used to measure WI and achieve water restricted conditions without altering the normal environmental and pen dynamics for the animals. Animal weights were measured at the beginning and the end of each trial (baseline and water restriction). Rumen samples were collected via rumenocentesis at the end of each trial. All samples were snap frozen on dry ice at sample collection, transported immediately to the lab and stored at $-80C^0$ until DNA extraction.

2.2.3. Animal selection for DNA extraction

Even though rumenocentesis was attempted on all animals, DNA extraction (for 16S rRNA gene amplicon sequencing) from all rumen samples is not cost effective. Hence, animals that showed statistically significant ($p < 0.1$) performance based on percentage recovery of average daily gain (ADG) were selected, and their rumenocentesis samples were used for whole genomic DNA extraction.

$$\% \text{ recovery of ADG} = \frac{\text{ADG during restriction trial}}{\text{ADG during WI trial}} \times 100$$

The % recovery of ADG distributions were calculated separately for each group and the animals that fall in the tails of the distribution were selected to maximize statistical power. A total of 146 animals from the seven groups were selected for fecal and rumen DNA extraction based on the criteria mentioned above. This gave rise to a total of 209 rumen samples that comprised of 115 samples from the water intake trial (baseline rumen samples) and 94 samples from the water restriction trial (restriction rumen samples).

2.2.4. Selection of DNA extraction kits

A pilot study was conducted to establish a high efficacy genomic DNA extraction methodology for the extraction of genomic DNA from liquid rumen samples obtained via rumenocentesis. Five randomly selected rumenocentesis samples (5 biological replicates for DNA extraction via each kit) were subjected to genomic DNA extraction using three different commercially available DNA extraction kits from Qiagen (QIAGEN

Sciences Inc., Germantown, MD). Namely; QIAamp® Fast DNA Stool Mini Kit (FSM), QIAamp® PowerFecal DNA Kit (PF), and QIAamp® PowerFecal Pro DNA Kit (PFP).

DNA extraction was conducted based on manufacturer's specifications with modifications to the bead beating step and the final DNA elution step (detailed in the following section). 250µl of rumen fluid was used as starting material. Bead beating was conducted using a BeadBug6® homogenizer (Benchmark Scientific, Inc. Edison, NJ). Molecular grade water was used as the negative control, and ZymoBIOMICS® Microbial Community Standard (Zymo Research, Irvine, CA) was used as the positive control. DNA quality and yield of each extraction (five replicate extractions for each of the three kits) was measured using NanoDrop® 1000 (ThermoFisher Scientific, Waltham, MA) spectrophotometer. Based on the initial DNA quality and yields obtained, an attempt was made to further optimize the mechanical lysis step by replacing the beads inside the Power Bead tubes with 1mm Zirconia beads and 0.1mm Zirconia beads. Accordingly, two more extractions were done using the same rumen samples for the two different sizes of Zirconia beads.

All DNA extractions (including negative and positive controls) were sent to Novogene Co., Ltd (CA, USA) for 16S rRNA gene amplicon sequencing of the V4 region. Resulting sequencing data was analyzed using DADA2 (Callahan et al., 2016) and visualized using MicrobiomeAnalyst (Dhariwal et al., 2017).

2.2.5. Whole genomic DNA extraction from rumen samples

Based on the results from the pilot study, QIAamp® PowerFecal Pro DNA Kit with standard Power Beads was used to extract genomic DNA from rumen samples according to manufacturer's specifications subjected to modifications in the bead beating and final elution steps. Bead beating was done using the BeadBug® 6 homogenizer (Benchmark Scientific, Inc. Edison, NJ) for two 2min iterations at an rpm of 4000 separated by a resting period of 5mins on ice (Smith, 2011; Lim et al., 2018). The BeadBug® program was set to 4 cycles of 30sec each. With a 30sec resting period.

Final elution was done using molecular grade water as two iterations of 100µl, each with a 5min incubation at room temperature. DNA was eluted by centrifuging at 14000 rpm for 1min. The second elution was concentrated using the CentriVap DNA concentrator (Labconco, Kansas City, MO) by evaporating it down to 50µl. The second elution was then added to the first elution to arrive at a volume of 150µl in the final DNA elution.

2.2.6. Library preparation and 16S V4 rRNA gene amplicon sequencing

The PCR amplification of extracted DNA and 16S amplicon sequencing was performed by Novogene Corporation Inc., Sacramento, CA. The V4 region of the 16S ribosomal RNA gene (~250bp) was amplified using the specific primer pair 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGAC TACHVGGGTWTCTAAT-3'), with the forward primer carrying the barcode sequence. The Phusion® High-Fidelity PCR Master Mix (New England Biolabs) was used to carry out all PCR reactions.

PCR products were purified by running them on a 2% agarose gel and extracting them using a Qiagen® Gel Extraction Kit (Qiagen, Germany). The sequencing library preparation was done using a NEBNext® UltraTM DNA Library Prep Kit for Illumina. It was quality tested using Qubit 3.0 fluorometer and quantified using Q-PCR. Subsequently, an Illumina HiSeq 2500 platform was utilized for the sequencing of the library, generating ~250bp paired-end raw sequence reads.

2.2.7. 16S V4 amplicon sequencing data analysis

The computational analysis of 16S V4 amplicon sequence data was conducted using DADA2 version 1.8, according to standard operation procedures for Illumina data (Callahan et al., 2016). All DADA2 codes were run using R-Studio version 1.1.4 on R version 3.6.1. Briefly; The quality scores of forward and reverse sequence reads were plotted to identify any drastic drops in sequence quality, specifically towards the ends of the sequences, and no such significant drops were observed. Hence, no specific truncation of the ends of sequences were done during the pre-processing step. However, linker primer sequences and any sequences that had ambiguous bases and an expected error greater than 2 were removed using the 'filterAndTrim' command in DADA2. Subsequently, the error rates for the amplicon sequence data were modeled using 'learnErrors' function in DADA2 and the sequence reads were then de-replicated using the function 'derepFastq'. The de-replicated sequences were then used for the inference of sequence variants for each sample using the core sample inference algorithms available in DADA2, implemented in the function 'dada'. Using the 'mergePairs' function, the corresponding forward and reverse inferred sequence variants were then merged to arrive at the full de-noised sequence contigs. The 'makeSequenceTable'

function was run on these contigs to create the amplicon sequence variant (ASV) table. The ASV table was then subjected to chimera removal and taxonomic classification via 'removeBimeraDenovo' and 'assignTaxonomy' functions. The DADA formatted version of the Silva version 132 reference database (Quast et al., 2013) was used as the training dataset for taxonomic classification. The ASV table and the taxonomy file generated by DADA2 was then used for statistical analysis and data visualization, using MicrobiomeAnalyst (Dhariwal et al., 2017).

2.2.8. Data filtering, normalization, statistical analysis and visualization

Low abundance features, defined as those with less than 10 read counts in at least 20% of the samples were removed prior to data normalization. Rarefying of data based on minimum library size, and Total Sum Scaling was performed as data normalization steps to negate the effect of uneven sequencing depth (Dhariwal et al., 2017). The resulting filtered and normalized data were used for subsequent statistical analyses and data visualizations. Baseline rumen samples from all selected animals were grouped together when making inferences about the rumen microbiome during *ad libitum* water intake and restriction rumen samples from all animals were grouped together when making inferences about the rumen microbiome during water restriction.

2.3. Results

2.3.1. Rarefaction analysis

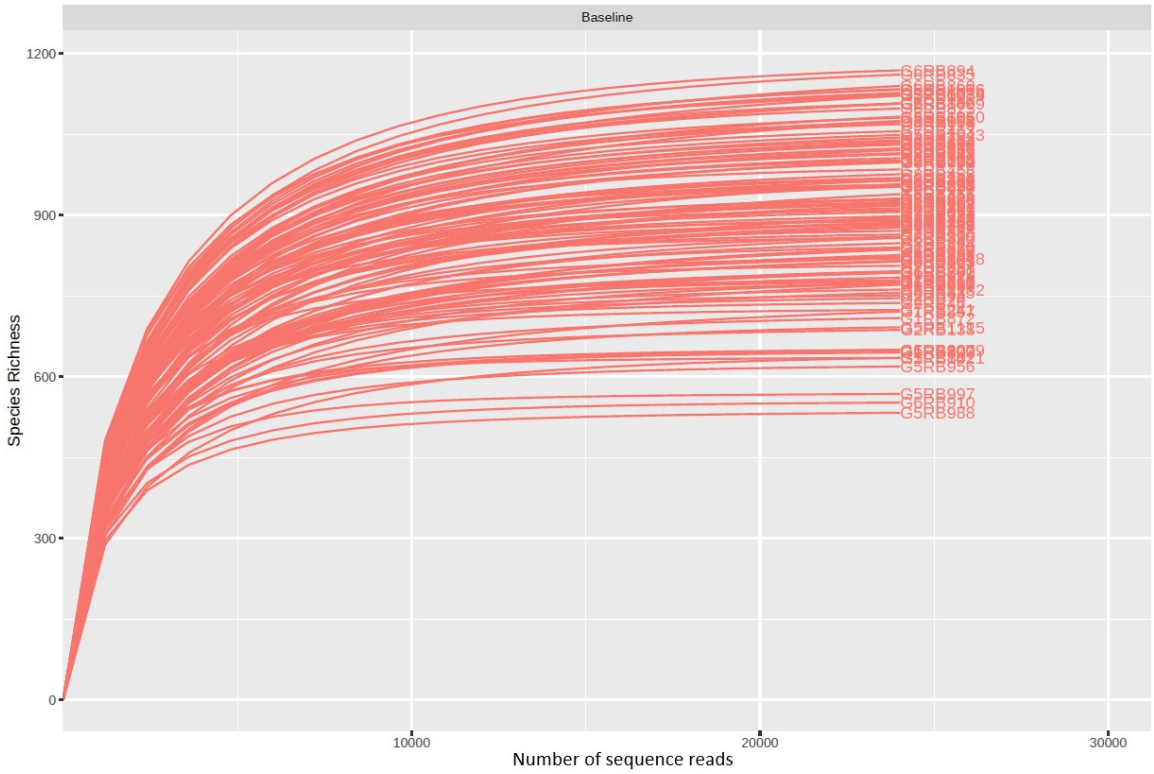
A total of 8,876,490 high quality sequences (94,431 ± 25,689) ranging from a minimum of 26,091 in a sample to a maximum of 126,820, were obtained from rumen restriction

samples. For rumen baseline samples, a total of 9,895,310 high quality sequences were obtained with a minimum sequencing read depth of 26,470 to a maximum read depth of 127,617 (86,046 ± 25,828).

To determine whether enough sampling depth was achieved to capture the prevalent microbial richness in baseline and restriction rumen samples during NGS of 16S V4 amplicons, a rarefaction analysis was performed generating rarefaction curves and Good's coverage values (Dethlefsen et al., 2008) for each rumen sample. Rarefaction analysis is a non-parametric re-sampling method that generates individual/sample based re-sampling species curves that are often used in microbiome studies to assess the completeness of the taxonomic survey in terms of identified taxa (Dethlefsen et al., 2008; McMurdie and Holmes, 2014).

The rarefaction curves generated for rumen samples originating from selected animals belonging to both WI (baseline) and water restriction trials are depicted in Figure 2.1. Plateaued curves were observed for all rumen samples, well before the count of re-sampled sequences equaled minimum library size (26,181 read counts). The plateauing of the curve after exponential increase at the beginning indicates the increasing sequencing effort required to identify new taxa (Dethlefsen et al., 2008). The plateaued curves for all rumen samples, subsequently becoming parallel to x-axis even before the minimum library size is reached, indicates that all rumen genomic DNA samples were sequenced with sufficient depth to capture the entirety of the microbial richness present in them.

Good's coverage on the other hand is a similar estimate based on the taxonomic classification of newly sampled reads. It determines the probability of the subsequently sampled read being classified into a taxon that has not been captured by the reads that were sampled before (Dethlefsen et al., 2008). For all rumen samples, the Good's coverage reached >99% at a sampling depth of ~24,000 sequence reads, meaning essentially that the total microbial diversity in the samples were captured at this point.



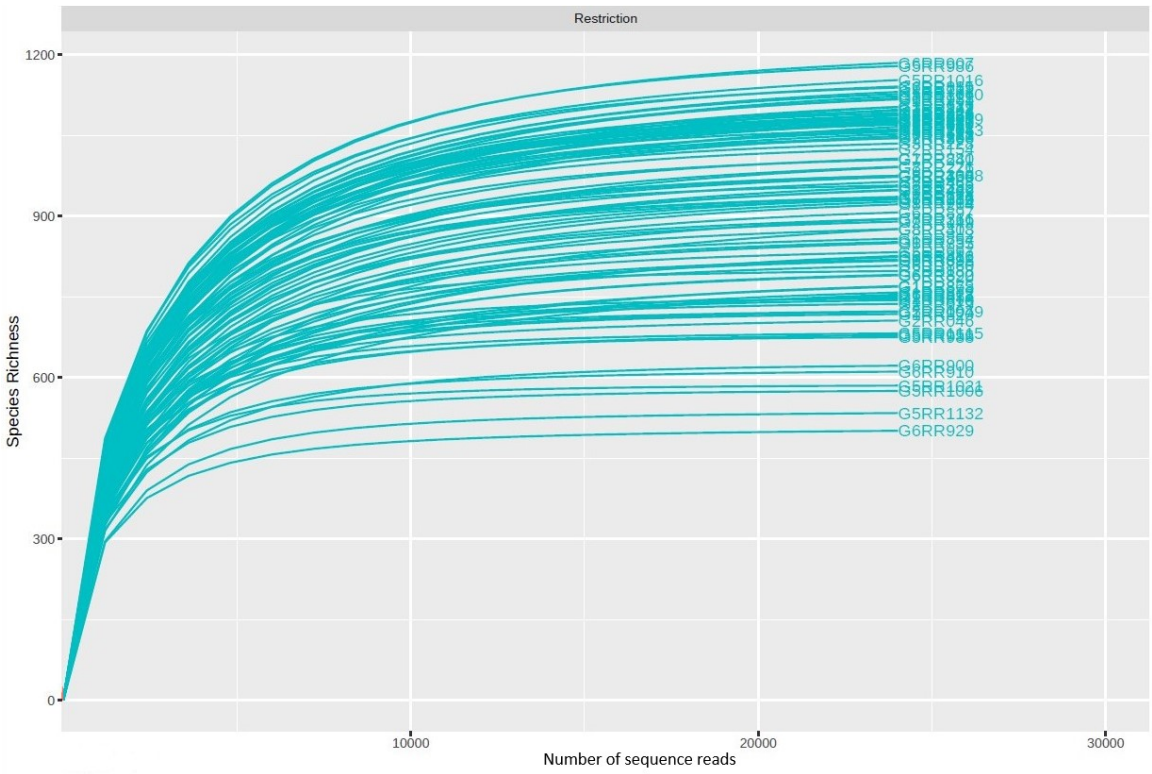


Figure 2.1: Rarefaction curves for baseline (top) and restriction (bottom) rumen samples. The curves were generated using the 'ggrare' function ranacapa R pckage (Kandlikar et al., 2018) implemented in MicrobiomeAnalyst (Chong et al., 2020), and demonstrates how the discovery of new species (species richness in y axis) increases with the number of sampled reads from each sample (x axis). All curves are truncated at a sampling depth of 26,181 sequence reads due to subsampling based on the minimum library size for all rumen samples (data rarefying).

2.3.2. Alpha diversity analysis

In order to determine the effect of water restriction on species richness and evenness of the rumen microbial population, we conducted an alpha diversity analysis on the microbial community abundance data recovered from baseline and restriction rumen DNA samples. Alpha diversity estimates such as Chao1 and Observed ASV indices were

used to calculate species richness (Ocejo et al. 2019), Shannon and Simpson indices are commonly used to measure alpha diversity of a microbial community based on both species evenness and richness (Wagner et al., 2018). Fisher's index models alpha diversity as a log series distribution and is seen as measure that takes species evenness into account (Parsons et al., 2017; Chong et al., 2020). Between group comparisons for Chao1, Observed ASV and Fisher's indices were performed using ANOVA, while Man-Whitney test was used for Shannon and Simpson alpha diversity estimates (Clemmons et al., 2017). The alpha diversity analysis was done using the phyloseq R package implemented in MicrobiomeAnalyst (McMurdie and Holmes, 2013; Chong et al., 2020).

No statistically significant differences ($p<0.05$) were observed in species richness and evenness (based on any of the alpha diversity indices mentioned above) between restriction and baseline samples. However, when subsampling (or data rarefying) was not used as a data normalization step (normalizing only via total sum of scaling or log ratio normalization) a statistically significant difference ($p<0.05$) was observed between restriction and baselines samples, based on Fisher's test for alpha diversity.

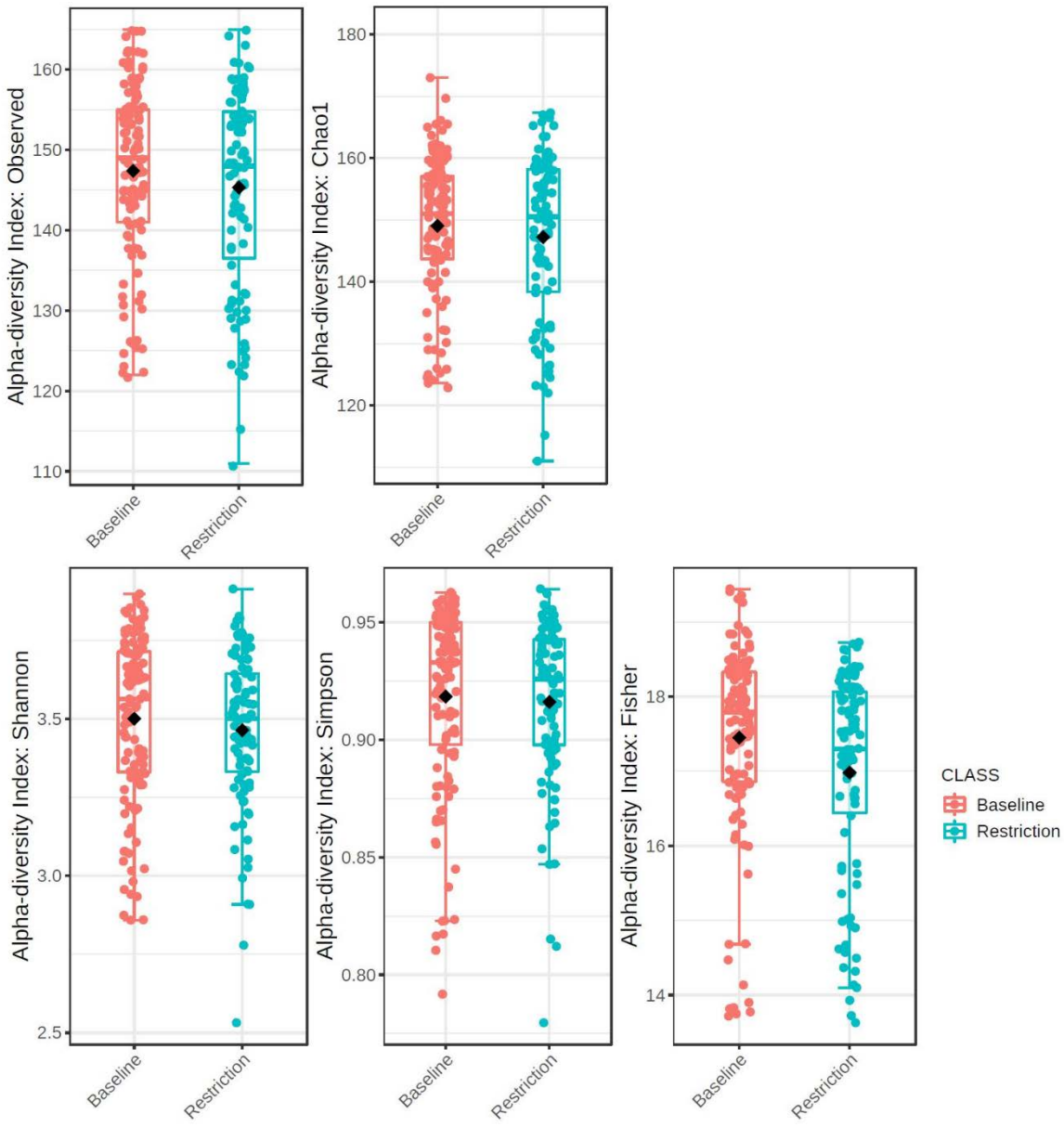


Figure 2.2: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from baseline and restriction rumen DNA samples. All rumen samples are grouped as either baseline or restriction and color coded accordingly. Chao1 and Observed ASV indices measure species richness while Shannon, Simpson, and Fisher indices measure species evenness.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	1.2499	0.21289
Chao1	T-test	1.0429	0.29837
Shannon	Man-Whitney	5976	0.18963
Simpson	Man-Whitney	5991	0.17825
Fisher's	T-test	2.4963	0.01338

Table 2.1: Between group comparisons of alpha diversity indices derived for baseline and restriction rumen samples. T-test was used to measure between group comparisons for the distribution of Observed ASV, Chao1, and Fisher's indices. Mann-Whitney test was used for Shannon and Simpson indices. Only Fisher's index indicated a statistically significant difference (in species evenness) when the data is normalized only using log ratio techniques or data scaling techniques (omitting subsampling from the data normalization process).

2.3.3. Beta diversity analysis

A beta diversity analysis was conducted based on the microbial community abundance data obtained from baseline and restriction rumen DNA samples to determine the effect of water restriction on the composition of rumen microbial community. Bray-Curtis dissimilarity and Jensen-Shannon index (Chong et al., 2020) were used to measure dissimilarity/similarity between baseline and restriction microbial communities based on the presence and absence of species. As depicted in Figure 2.3, when visualized using NMDS (None-parametric multidimensional scaling) (Luz Calle, 2019) plots, the baseline and restriction samples did not form clear visible clusters. However, when the overall clustering pattern was analyzed using permutational analysis of variance (PERMANOVA) (Moore et al., 2017), a statistically significant ($p<0.05$) difference in

species composition was observed between the microbial communities recovered from baseline and restriction rumen samples for both the dissimilarity measures used. The statistics for Bray-Curtis dissimilarity were; F-value: 2.5244; R-squared: 0.012048; p-value <0.03; [NMDS] Stress = 0.18029. And the statistics for Jensen-Shannon index were; F-value: 3.3465; R-squared: 0.01591; p-value <0.031; [NMDS] Stress = 0.17346.

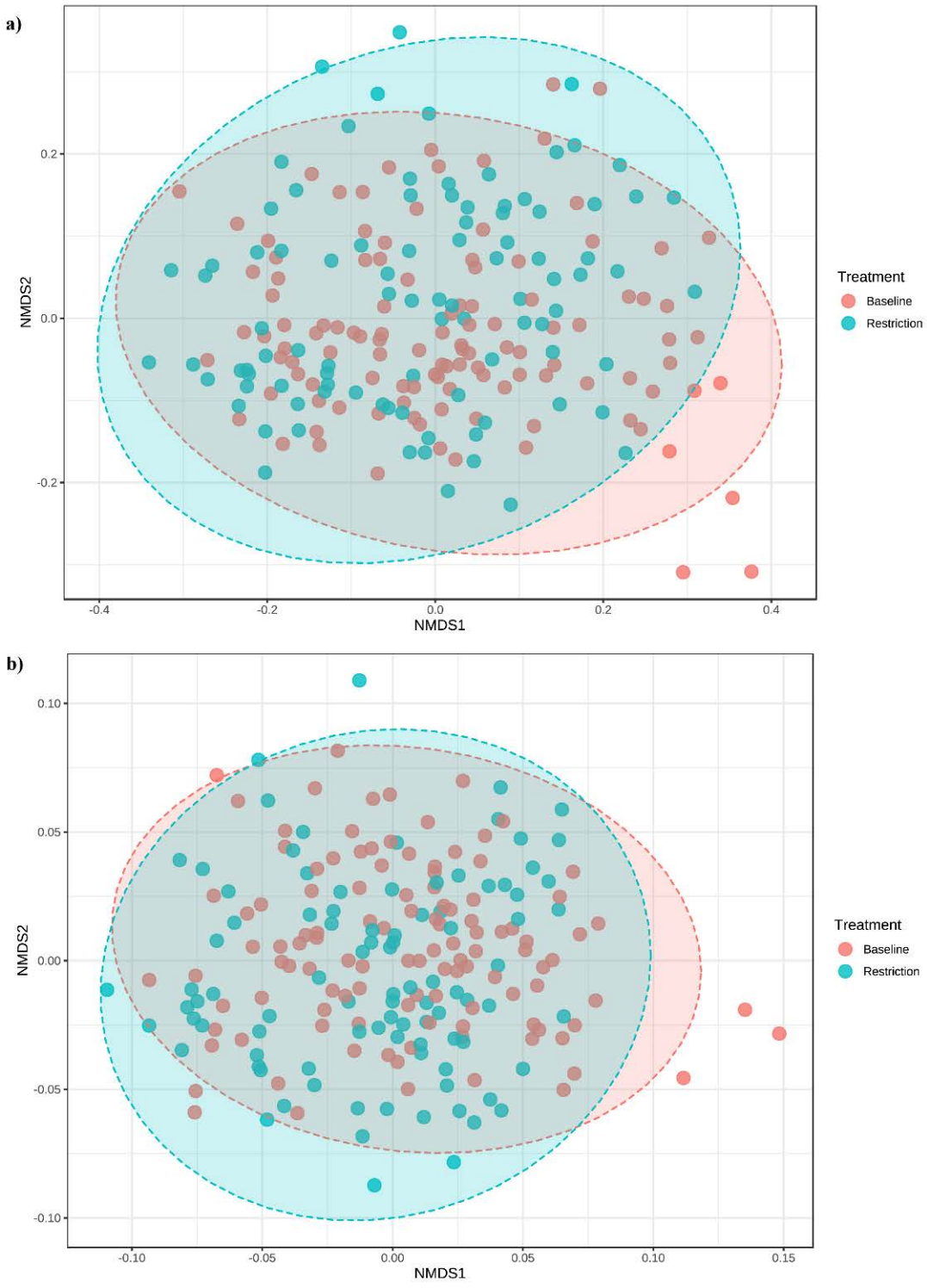


Figure 2.3: NMDS plots illustrating dissimilarity matrices generated using Bray-Curtis (a) and Jensen-Shannon (b) indices, in the two-dimensional space. Each blue (restriction rumen samples) and red (baseline rumen samples) point represent the species

composition of the entire microbiome present in a given rumen sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups. Plot (a) NMDS stress = 0.18029. Plot (b) NMDS stress = 0.17346.

2.3.4. Microbial community composition

In order to assess the effect of water restriction on the rumen microbial community composition, taxonomic abundance profiles were built at phylum and genus levels using MicrobiomeAnalyst. A total of 1403 features (ASVs) that remained after data filtering and normalization was used to identify the 10 most abundant phyla and genera in baseline and restriction rumen samples.

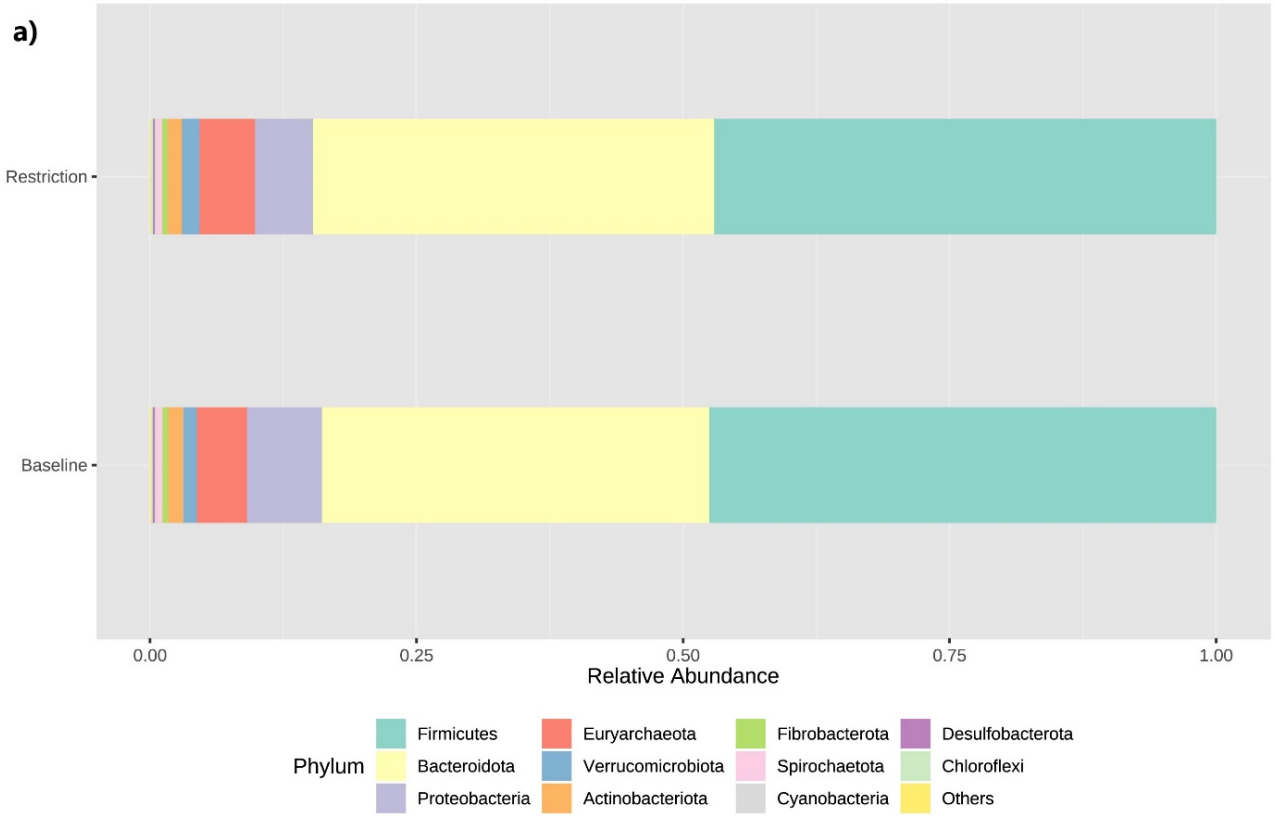
For baseline samples, out of the top 10 phyla, Firmicutes (47.5%), Bacteroidota (36.3%), Proteobacteria (7.0%), Euryarchaeota (4.7%), Actinobacteriota (1.4%), and Verrucomicrobiota (1.3%) accounted for 98.3% of the total abundance. The rest of the 10 most abundant phyla were Fibrobacterota, Spirochaetota, Cyanobacteria and Desulfobacterota, accounting for 1.4% of the total abundance. For restriction samples, the top 10 and the top 6 phyla (accounting for 98.4% of the total abundance) essentially remained the same.

However, a statistically significant drop in relative abundance was observed for phyla Proteobacteria (baseline 7.01% to restriction 5.40%; $p < 0.05$) and Desulfobacterota (baseline 0.18% to restriction 0.15%; $p < 0.05$) in the restriction samples. Furthermore, a significant increase in the relative abundance of phyla Euryarchaeota (baseline 4.72% to

restriction 5.24%; $p < 0.05$) and Verrucomicrobiota (baseline 1.26% to restriction 1.62%; $p < 0.01$) was observed in the restriction samples as well (Table 2.2).

As depicted in Figure 2.4(b), when the relative abundance data was summarized at genus level, the baseline samples were dominated by genera such as *Prevotella* (21.9%), *Christensenellaceae_R_7_group* (6.4%), *Lachnospiraceae_NK3A20_group* (6.3%), *Methanobrevibacter* (4.6%), *Succinivibrionaceae_UCG_002* (3.6%), and *Ruminococcus* (3.3%). These top 6 genera accounted for a cumulative 46.1% of the total microbial abundance. The other four genera out of the top 10 were *NK4A214_group, Rikenellaceae_RC9_gut_group, Lachnospiraceae_XPB1014_group,* and *Prevotellaceae_UCG_003*, giving rise to a cumulative 7.3% of the total microbial abundance. 20.1% of the total abundance could not be assigned a classification at genus level.

The taxonomy of the 10 most predominant genera in restriction rumen samples, remained the same as the baseline samples as depicted in Figure 2.4(b). However, a significant increase in the relative abundance of *Methanobrevibacter* (baseline 4.6% to restriction 5.1%; $p < 0.05$) and *Rikenellaceae_RC9_gut_group* (baseline 1.95% to restriction 2.14%; $p < 0.05$) was observed in the restriction samples while the abundance of *Prevotellaceae_UCG_003* decreased significantly (baseline 1.61% to restriction 1.47%; $p < 0.05$). Marginally significant ($p<0.05$) variations in relative abundance was observed for *Succinivibrionaceae_UCG_002* and *NK4A214_group* as well.


a)
Restriction
Baseline
0.00
0.25
0.50
0.75
1.00
Relative Abundance
Phylum
Firmicutes
Bacteroidota
Proteobacteria
Euryarchaeota
Verrucomicrobiota
Actinobacteriota
Fibrobacterota
Spirochaetota
Cyanobacteria
Desulfobacterota
Chloroflexi
Others

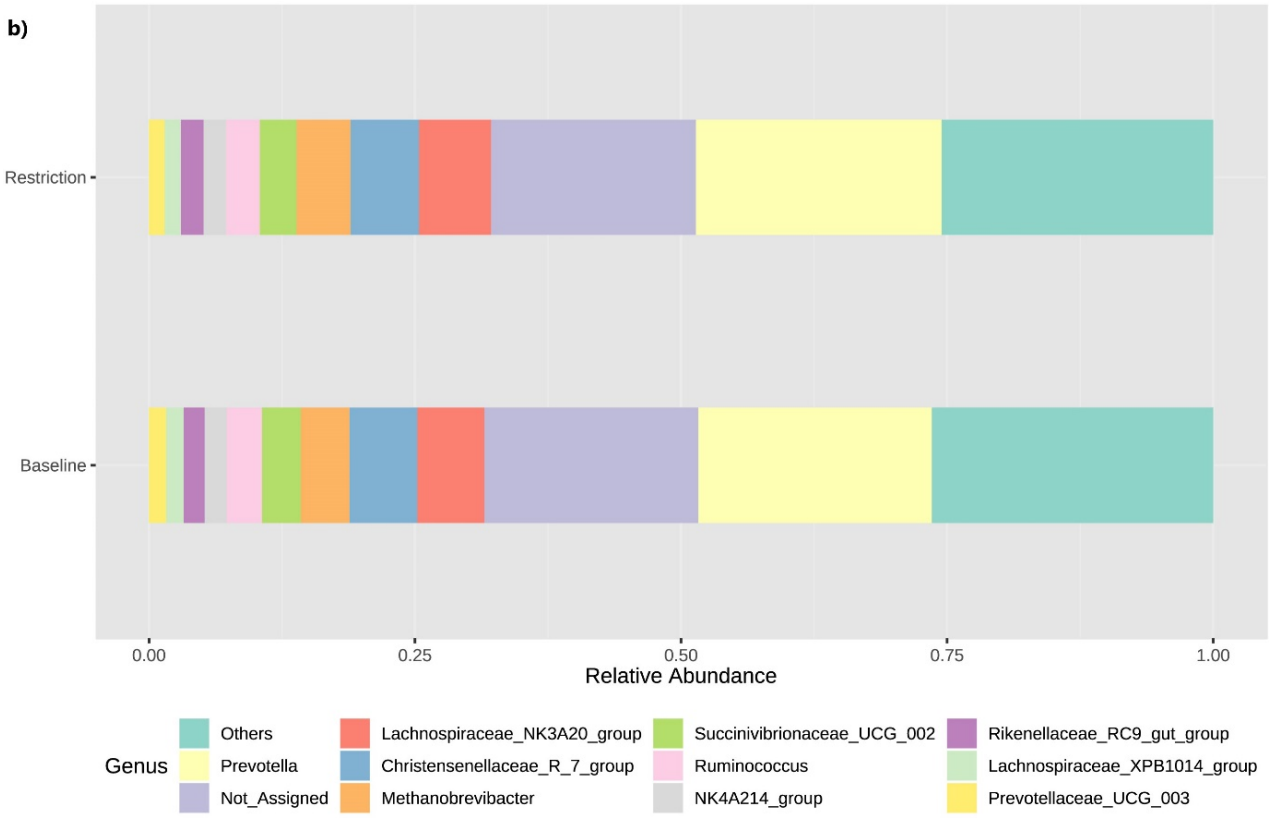

b)
Restriction
Baseline
0.00
0.25
0.50
0.75
1.00
Relative Abundance
Genus
Others
Prevotella
Not_Assigned
Lachnospiraceae_NK3A20_group
Christensenellaceae_R_7_group
Methanobrevibacter
Succinivibrionaceae_UCG_002
Ruminococcus
NK4A214_group
Rikenellaceae_RC9_gut_group
Lachnospiraceae_XPB1014_group
Prevotellaceae_UCG_003

Figure 2.4: Relative abundance bar graphs depicting the 10 most predominant phyla (a) and genera (b) present in baseline and restriction rumen samples.

Phyla	Baseline	Restriction	P-value
Firmicutes	47.54%	47.08%	0.7611
Bacteroidota	36.35%	37.66%	0.1857
Proteobacteria	7.01%	5.40%	0.0002***
Euryarchaeota	4.72%	5.24%	0.0018**
Verrucomicrobiota	1.26%	1.62%	0.0011**
Actinobacteriota	1.43%	1.36%	0.8499
Fibrobacterota	0.54%	0.52%	0.1461
Spirochaetota	0.52%	0.47%	0.0559
Cyanobacteria	0.19%	0.22%	0.4993
Desulfobacterota	0.18%	0.15%	0.0414*

(*) – p <0.05 (**) – p <0.01 (***) – p<0.001

Table 2.2: The 10 most abundant phyla observed in the baseline and restriction rumen samples of cattle. Baseline and Restriction columns indicate the relative abundance of each phylum in baseline and restriction rumen samples. The p-values indicated are between group (baseline and restriction) comparisons for each phyla using paired t-test.

Genus	Baseline	Restriction	P-value
Prevotella	21.92%	23.09%	0.250
Lachnospiraceae_NK3A20_group	6.30%	6.77%	0.054
Christensenellaceae_R_7_group	6.38%	6.45%	0.220
Methanobrevibacter	4.57%	5.06%	0.001**
Succinivibrionaceae_UCG_002	3.64%	3.44%	0.041*
Ruminococcus	3.31%	3.16%	0.228
NK4A214_group	2.09%	2.13%	0.049*
Rikenellaceae_RC9_gut_group	1.95%	2.14%	0.006**
Lachnospiraceae_XPB1014_group	1.66%	1.52%	0.150
Prevotellaceae_UCG_003	1.61%	1.47%	0.019*

(**) – p < 0.01 (*) – p < 0.05

Table 2. 3: The relative abundance of the 10 most predominant genera observed in the rumen samples obtained during the water intake (baseline) and water restriction

(restriction) trials. The p-values were calculated using paired t-test on animal wise abundant data grouped according to treatment.

2.3.5. The core-microbiome

To further elucidate the microbial community structure of baseline and restriction rumen samples, we conducted a core-microbiome analysis at genus level. The core microbiome was defined as the taxa that are present in at least 70% of the samples considered, at a threshold relative abundance of 0.01% or above. The core-microbiome for baseline rumen samples (as depicted in Figure 2.5a) consisted of 9 genera, namely; *Ruminococcus, Prevotella, Lachnospiraceae_NK3A20_group, Christensenellaceae_R_7_group, Methanobrevibacter, Oscillospiraceae_NK4A214_group, Rikenellaceae_RC9_gut_group, Prevotellaceae_UCG_003,* and *Mogibacterium.*

The core microbiome for rumen samples from the water restriction trial (as depicted in Figure 2.5b) included 10 genera comprising of *Ruminococcus, Prevotella, Lachnospiraceae_NK3A20_group, Christensenellaceae_R_7_group, Methanobrevibacter, Oscillospiraceae_NK4A214_group, Rikenellaceae_RC9_gut_group, Prevotellaceae_UCG_003, Lachnospiraceae_XPB1014_group* and *Mogibacterium.*

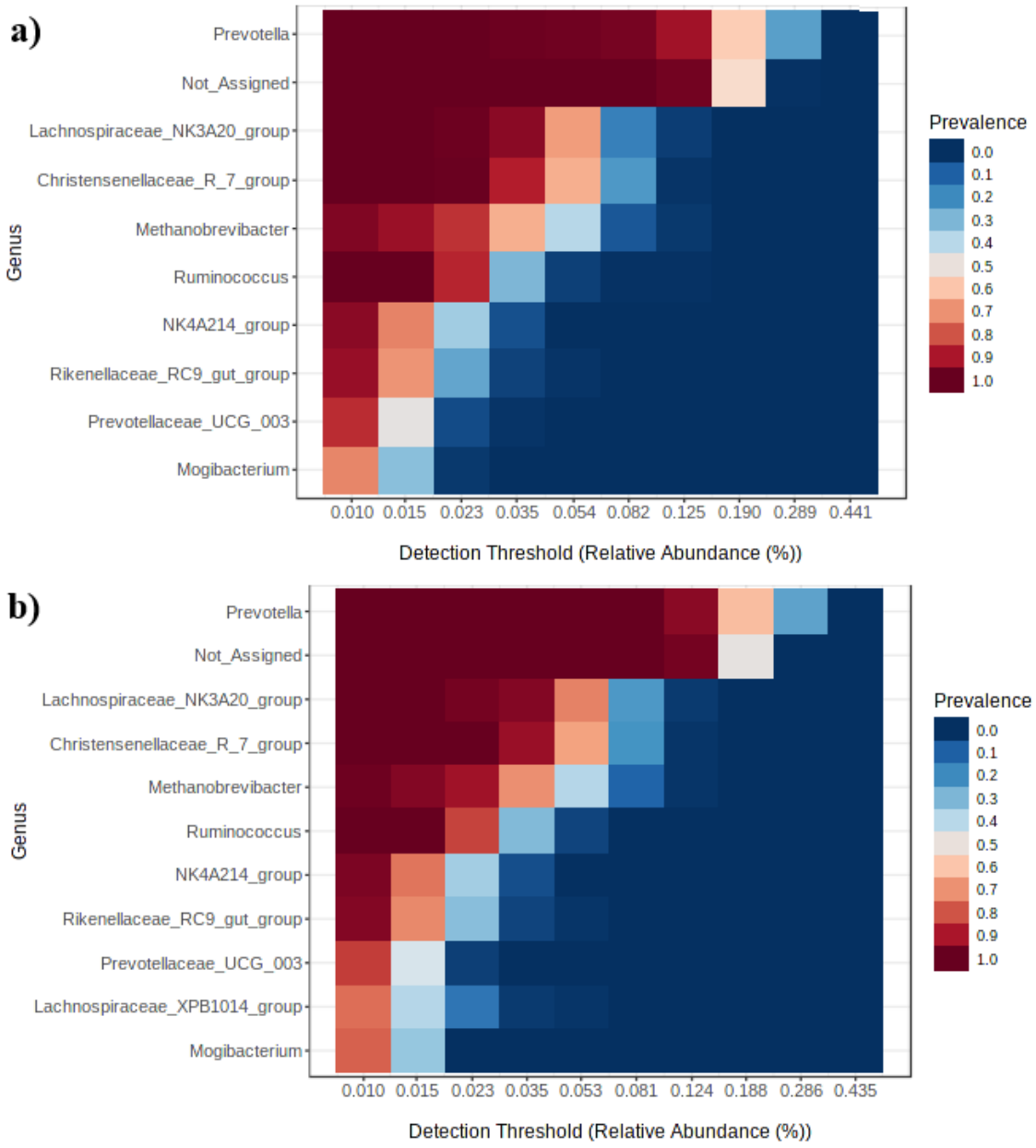


Figure 2.5: The heat map illustrating the relative abundance and prevalence of the microbial taxa defining the core-microbiomes recovered from baseline (a) and restriction (b) rumen samples. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria, for microbial taxa to be considered as a part of the baseline or restriction rumen core-microbiome.

2.3.6 Biomarker analysis of baseline and restriction rumen microbial populations

In an effort to identify differentially abundant and biologically significant taxa, or biomarkers, enriched in baseline and restriction rumen samples, a linear discriminant analysis (LDA), also known as LEfSe (LDA effect size), was conducted (Segata et al., 2011). The LEfSe algorithm takes into account both statistically significant differential abundance of taxa and the magnitude of the effect each of these taxa would have on classifying samples into treatment groups (also known as biological consistency). Kruskal-Wallis rank-sum test is used to identify statistically significant taxa and LDA is used to measure their effect size (Segata et al., 2011). Data summarized at genus level was used for this analysis, and a selection criteria of $p<0.05$ and LDA > 2.0 was used to identify differentially abundant taxa (biomarkers) associated with baseline and restriction rumen samples.

We were able to identify a total of 35 differentially abundant taxa as biomarkers using the LEfSe analysis. The 15 most predominant biomarkers out of that 35, are depicted in Figure 2.6. Based on that, it can be observed that *UCG-004 (Erysipelatoclostridiaceae* family*), Prevotellaceae_UCG_004, Moryella, Saccharofermentans, Selenomonas, Flexilinea, Lachnoclostridium* were enriched in restriction rumen samples. *Bacteroides, CAG_352 (Ruminococcaceae* family*), Escherichia_Shigella, Romboutsia, UCG-005* (*Oscillospiraceae* family), *Ruminobacter, Turicibacter* and *Succinivibrionaceae_UCG_001* were identified as enriched biomarkers in baseline rumen samples.

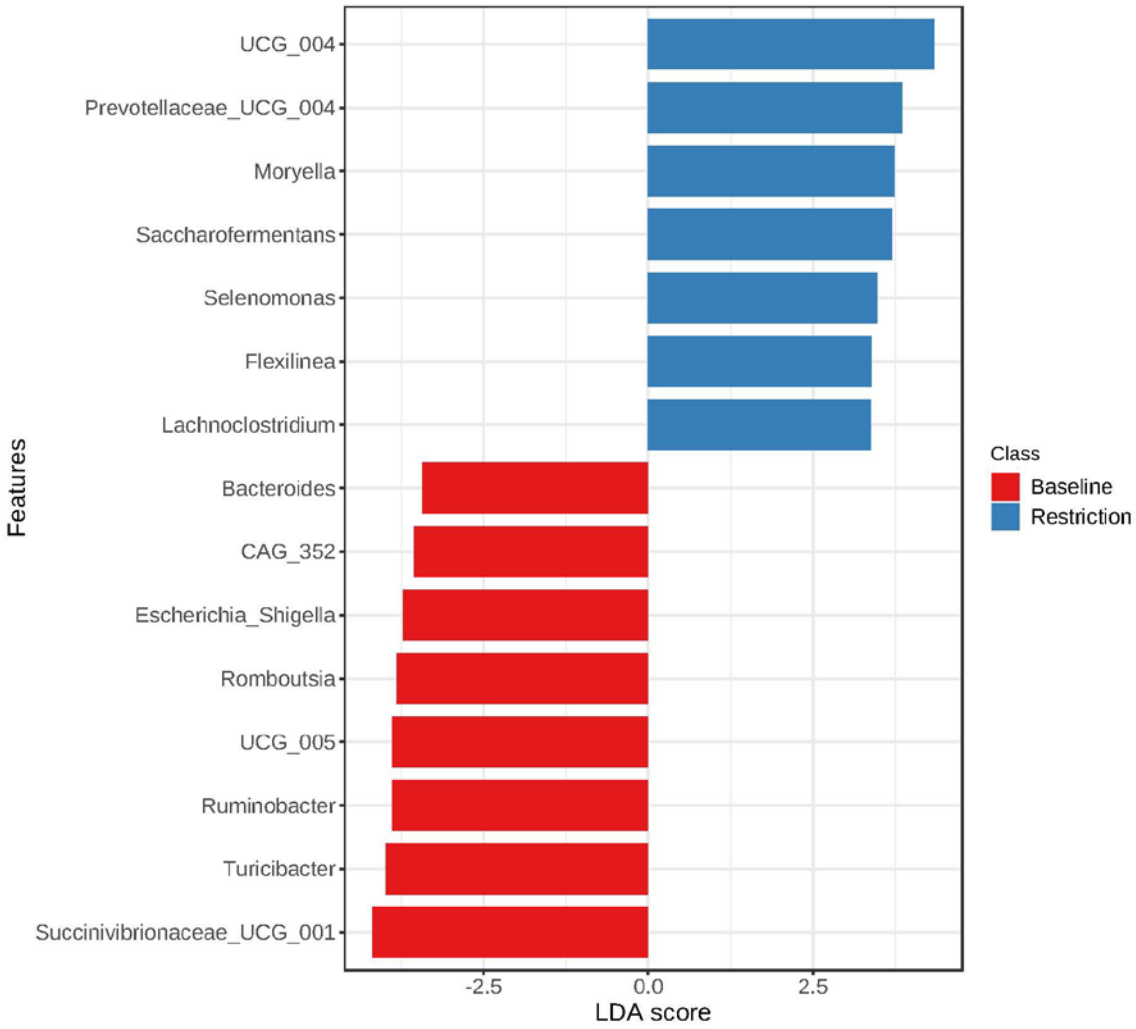


Figure 2.6: The bar graph depicting the results of linear discriminant analysis (LDA) effect size (LEfSe) done on baseline and restriction rumen samples. The blue bars indicate LDA score of biomarkers for rumen restriction samples and the red bars indicate LDA scores of biomarkers for rumen baseline samples. The LDA score is an estimation of the magnitude of the effect each biomarker would have in grouping each sample into the two classes (baseline and restriction) (Segata et al., 2011; Chong et al., 2020).

Random forest (RF) analysis (Breiman, 2001) implemented in MicrobiomeAnalyst (Chong et al., 2020) was used to further elucidate the differentially abundant taxa between baseline and restriction rumen samples. The Random Forest (RF) algorithm, originally developed by Leo Breiman (Breiman, 2001), is a supervised learning algorithm that can be well adapted to analyze high dimensional microbiome data (Chong et al., 2020). Due to its ability to efficiently analyze large datasets, accurately classify features into treatment groups, and determine the importance of each feature for classification (as a proxy for biological consistency), it is utilized in microbiome studies to discover enriched features in different environments or treatment groups.

The out-of-sample error (how accurately the algorithm can classify new data) also known as the OOB error for the RF algorithm converges to a minimum (limit) when the number of decision trees built, increases (Breiman, 2001). Hence, we generated 5000 decision trees (the maximum allowed by MicrobiomeAnalyst) when looking for enriched taxa in baseline and restriction rumen samples, thus lowering the OOB error to 0.263. Lower OOB errors are associate better sensitivity and classification accuracy of underlying random forest generated (Roguet et al., 2018).

Out of the top 15 biomarkers identified by RF analysis (depicted in Figure: 2.7), 13 overlapped with the taxa identified by LEfSe and the rest, namely; *Catenibacterium, Saccharopolyspora, Geobacillus* were unique to RF analysis.

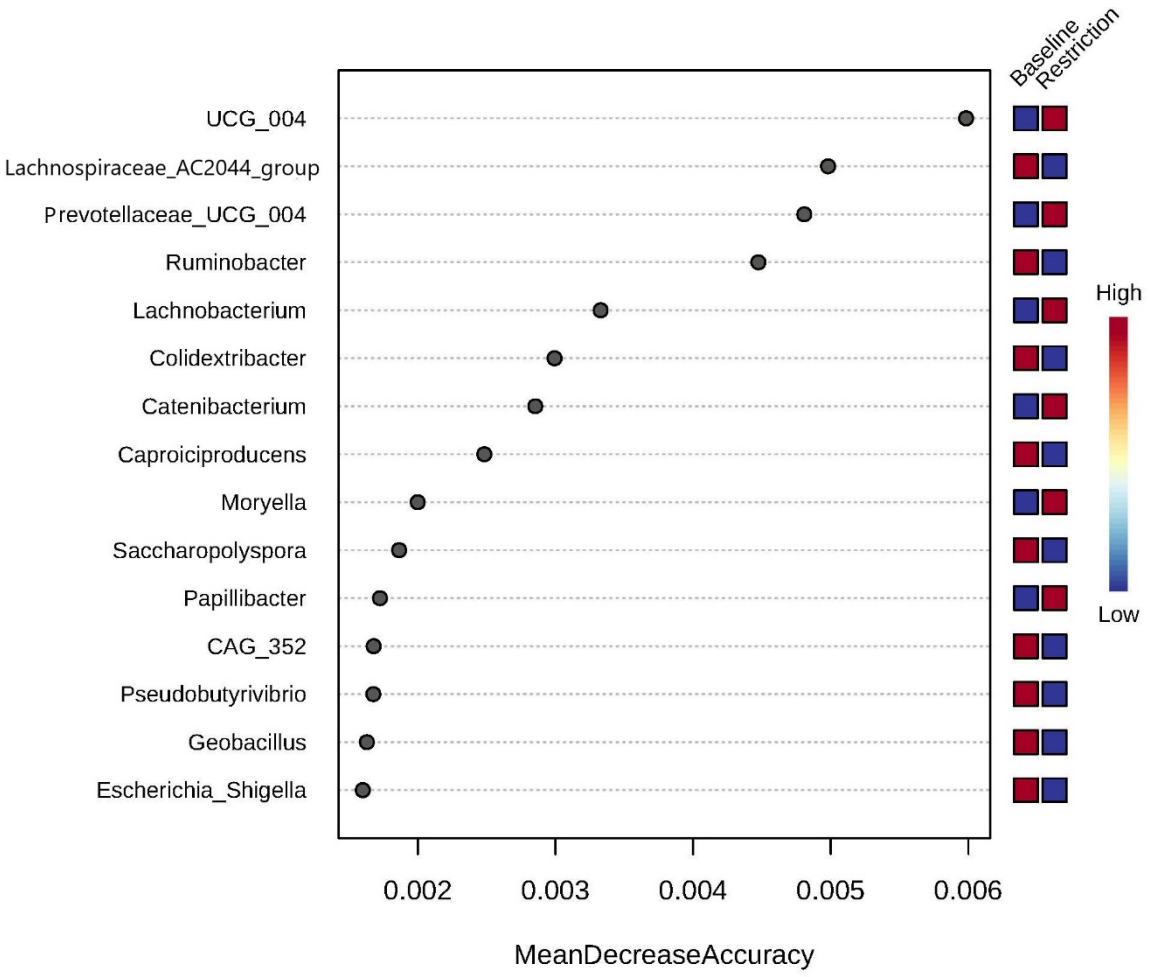


Figure 2.7: Illustration of the differentially abundant features identified by random forest analysis. The features are ranked according to the mean decrease in accuracy the model suffers in assigning samples to the classes, if the particular feature is removed from the analysis. The color codes indicate the association of each feature either with baseline or restriction samples.

Genus	P-values	FDR	LDAscore	Association
Succinivibrionaceae_UCG_001	4.672e-02	1.800e-01	-4.19	Baseline
Turicibacter	3.367e-02	1.641e-01	-3.99	Baseline
*Ruminobacter**	1.759e-02	1.324e-01	-3.88	Baseline
UCG_005	3.820e-02	1.659e-01	-3.88	Baseline
Romboutsia	4.913e-02	1.813e-01	-3.82	Baseline
*Escherichia_Shigella**	2.242e-03	4.196e-02	-3.72	Baseline
*CAG_352**	2.417e-02	1.439e-01	-3.55	Baseline
Bacteroides	2.127e-02	1.327e-01	-3.43	Baseline
Cellulosilyticum	2.809e-02	1.508e-01	-3.35	Baseline
Clostridioides	3.926e-02	1.659e-01	-3.34	Baseline
*Pseudobutyrivibrio**	3.873e-02	1.659e-01	-3.26	Baseline
*Lachnospiraceae_AC2044_group**	1.328e-05	8.700e-04	-3.2	Baseline
Alistipes	1.204e-02	1.052e-01	-3.18	Baseline
Tuzzerella	6.065e-03	6.621e-02	-3.1	Baseline
*Caproiciproducens**	5.054e-04	1.655e-02	-2.92	Baseline
Catenisphaera	4.416e-02	1.753e-01	-2.66	Baseline
Flavonifractor	3.383e-02	1.641e-01	-2.65	Baseline
Oscillibacter	2.021e-02	1.324e-01	-2.56	Baseline
*Colidextribacter**	5.275e-03	6.621e-02	-2.53	Baseline
Parabacteroides	3.582e-02	1.659e-01	-2.47	Baseline
Parvibacter	1.536e-02	1.257e-01	-2.45	Baseline
*UCG_004**	1.722e-06	2.255e-04	4.34	Restriction
*Prevotellaceae_UCG_004**	2.870e-05	1.253e-03	3.86	Restriction
*Moryella**	4.485e-03	6.528e-02	3.73	Restriction
Saccharofermentans	2.877e-02	1.508e-01	3.7	Restriction
Selenomonas	1.106e-02	1.052e-01	3.48	Restriction
Flexilinea	3.787e-03	6.201e-02	3.38	Restriction
Lachnoclostridium	1.925e-02	1.324e-01	3.37	Restriction
Oribacterium	2.743e-02	1.508e-01	3.22	Restriction
*Lachnobacterium**	1.108e-03	2.903e-02	3.21	Restriction
Veillonellaceae_UCG_001	5.929e-03	6.621e-02	3.19	Restriction
*Papillibacter**	1.450e-03	3.165e-02	3.08	Restriction
Anaerovorax	1.139e-02	1.052e-01	2.84	Restriction
Defluviitaleaceae_UCG_011	4.073e-02	1.667e-01	2.77	Restriction
Lachnospiraceae_UCG_009	1.916e-02	1.324e-01	2.16	Restriction

Table 2.4: The 35 differentially abundant taxa (biomarkers) identified using LEfSe analysis, and their overlap (*) with the top 15 most predominant biomarkers identified by RF analysis.

2.4. Discussion

Ruminants are often considered to be superorganisms due to the prominent symbiotic host-microbiome relationship they maintain with the rumen microbiota (Matthews et al., 2019). The rumen is home to a diverse and complex microbial community that comprises of anaerobic bacteria, methanogenic archaea, fungi, protozoa and phages, and facilitates pre-gastric fermentation in ruminants (Weimer et al., 2009; Huws et al., 2018).

The diversity and abundance of rumen microbiota can be affected by a multitude of internal (host) and external environmental factors. Internal factors such as host genetics (Roehe et al., 2016; Li et al., 2019; Fan et al., 2020), age (Liu et al., 2017), sex (Li et al., 2019), and external factors such as feed intake (Delgado *et al.*, 2019), diet (Pitta *et al.*, 2010), feed withdrawal (Rabaza et al., 2019), and heat stress (Tajima et al., 2007; Romero-Pérez et al., 2011; Zhao et al., 2019; Baek et al., 2020) have been associated with changes in the rumen microbial population via seminal and more recent studies found in existing literature. However, substantial studies on the direct and indirect effects of water restriction on the rumen microbial community are yet to materialize.

In order to assess the effect of water restriction on the rumen microbial dynamics, we reconstructed microbial communities in rumenocentesis samples obtained from animals

maintained at 50% water restriction, using 16S rRNA V4 gene amplicon sequencing and compared them against similarly reconstructed microbial communities from the baseline rumenocentesis samples of the same animals.

2.4.1 Alpha diversity measures indicate a possible difference in species evenness between restriction and baseline rumen microbial communities.

Alpha diversity, also known as within sample diversity, is a measure of species richness (the number of different species available in a given environment) and species evenness (the consistency of species abundance). Diversity indices such as observed richness (observed ASV/OTU) and Chao1 are commonly used to assess the alpha diversity of a population based on species richness (Luz Calle, 2019; Chong et al., 2020). They do not take species evenness into account. Observed richness is often criticized for under representation of the actual richness in microbial community since under this method species with low abundance can go undetected. Hence, extended richness measures such as Chao1 that explicitly takes into account the number of species that are found once or twice, in its calculation are also utilized in parallel to assess alpha diversity of a microbial community in terms of species richness (Luz Calle, 2019).

Even though species richness indices calculated (Chao1 and observed ASV) for baseline samples were slightly higher than those that were for restriction rumen samples, they did not exhibit significant differences ($p<0.05$) in their pairwise comparisons. Thus, making it clear that, overall, there was no significant difference in the number of different species

in the microbial communities recovered from the rumen samples when the animals were under water restriction versus *ad libitum* water intake.

Similar to species richness indices, Shannon and Simpson indices did not exhibit a statistically significant difference ($p<0.05$) during pairwise comparison (Man-Whitney test), even though baseline values were slightly higher than restriction values. The Fisher's index however did show a significant difference in evenness between baseline and restriction samples, when the data was not rarefied/sub-sampled (all other indices revealed statistically insignificant pairwise comparison results with or without data rarefying). This may well be due to the fact that, even though data rarefying based on minimum library size is a popular method for normalizing data from samples with vastly different sequence library sizes, there is an apparent loss of information since a significant amount of sequence reads are discarded from samples with large sequencing libraries, thus affecting alpha diversity estimations (McMurdie and Holmes, 2014; Willis, 2019).

Our observations indicate that water restriction may possibly be exerting an effect on the relative abundance of individual species in the rumen microbiome rather than affecting the total number of unique species that comprises it. Furthermore, this may be a direct effect of water scarcity on the growth of certain ruminant microorganisms, or an indirect effect of water restriction, through its effect on feed intake and digestibility in cattle (Benatallah et al. 2019; Hackmann and Firkins 2015).

2.4.2 The species composition of rumen microbial communities recovered from baseline and restriction rumen samples are significantly different.

Dissimilarity matrices constructed using Bray-Curtis index, Jensen-Shannon divergence and Jaccard index for baseline and restriction rumen samples at genus level, did not portray clear clustering when visualized using NMDS plots (Ramette, 2007; Luz Calle, 2019). However, when their overall beta diversity was analyzed using PERMANOVA (Anderson, 2001; Moore et al., 2017; Luz Calle, 2019), the results were statistically significant ($p<0.05$), suggesting a significant difference in species composition in the microbial communities reconstructed from baseline and restriction rumen samples.

The reason for observing no clear clustering in the NMDS plots for statistically significant data, may be due to the information loss during dimensionality reduction. Microbiome data are high dimensional and the NMDS plots try to bring these down to 2 or 3 dimensions. Further evidence for this 'information loss' can be found in the stress values calculated for each plot which are close to 0.2, meaning that the NMDS plot may not be fully accurate in depicting the actual distributions of the samples in two dimensional space (Chong et al., 2020).

2.4.3 The taxonomy of the predominant rumen microbiota remains largely unchanged while their relative abundances differ.

Unlike observational data from any other natural ecosystem, high throughput sequencing (HTS) data from a particular microbiome is always limited by the sequencing capacity of

the HTS platform. The main implication of this phenomenon is that the abundance of a particular species cannot be considered as independent of the abundance of the other species co-inhabiting the environment. Hence taxonomic abundance data is best portrayed as relative abundances rather than actual abundances or count data (Gloor et al., 2017). Therefore, we used relative abundance graphs to summarize the taxonomic abundances obtained through sequence data analysis.

The predominant taxa observed by us in baseline and restriction rumen samples fall well in line with existing literature describing the taxonomic composition of the rumen microbiome (Seshadri et al. 2018; Henderson et al. 2015; Matthews et al. 2019; Snelling et al. 2019). The taxonomy of the 10 most abundant genera and phyla remained the same between baseline and restriction samples, however we observed significant variations in their relative abundances, both at phylum and genus level.

A significant drop ($p<0.05$) in the relative abundance of phylum Proteobacteria observed in restriction samples was particularly noteworthy. This was mainly caused by a drastic drop in the abundance of *Ruminobacter* (45% reduction), *Succinivibrio* (13% reduction), *Escherichia/Shigella* (78% reduction) and *Acinetobacter* (30% reduction) at genus level in restriction rumen samples compared to baseline samples.

The genus *Ruminobacter* is known to contain major amylolytic bacteria found in the rumen responsible for starch degradation. They are also known to be highly dependent on starch as a substrate, hence observed to decrease in abundance when the animal is fed

with a sugar-rich diet (Klevenhusen et al., 2017). Abundance of *Ruminobacter* is known to increase in heat stressed animals under increased ruminal pH (Zhao et al., 2019). Animals used for the current study did show a slight but a significant decrease (p=8.29e-05) in rumen pH, where it declined from an average baseline pH of 6.45 to 6.36 in restriction rumen samples. The behavior of *Ruminobacter* species, when subjected to water scarcity has not been described yet. The significant decrease in the abundance of *Ruminobacter* we observed during water restriction could be an indication of reduced starch digestibility of the rumen under water restricted conditions.

Escherichia/Shigella are mucosa-associated microorganisms, widely known for their adverse gut health implications in ruminants and humans, including biofilm formation during frothy bloat and zoonotic transmission of virulent strains such as *Escherichia. Coli* O157:H7 (Auffret et al. 2020; Pitta et al. 2016; Stein and Katz 2017). A 78% reduction of the baseline *Escherichia/Shigella* abundance during water restriction, can be considered as a beneficial health implication of water restriction. However, this requires further scrutinization of the data at species and strain level, since *Escherichia spp.* may also play an integral role in the natural symbiosis of the rumen microbiome.

The relative abundances of *Methanobrevibacter* and *Rikenellaceae_RC9_gut_group* showed a significant increase in their relative abundances in restriction rumen samples when compared to baseline samples. According to the global rumen census data (Matthews et al., 2019), *Methanobrevibacter* is most predominant methanogen found in ruminants. The abundance of *Methanobrevibacter* (or methanogens in general) in the

beef cattle rumen is known to differ based on factors such as residual feed intake and diet type (Carberry et al., 2014). However, current literature suggests a knowledge gap in assessing the effect of the water consumption on enteric methane production in ruminants and microbial dynamics of methanogens, including *Methanobrevibacter* (Hill et al., 2016). Hence, our study provides a first look at how predominant methanogens such as *Methanobrevibacter* may increase in abundance during water restricted conditions. However, whether this is a direct effect of water intake on the rumen methanogens or an indirect effect through differed feed intake or microbial co-occurrence networks, remains to be further investigated.

Rikenellaceae_RC9_gut_group extends from the family *Rikenellaceae* known to produce volatile fatty acids during the rumen fermentation process (Holman and Gzyl, 2019). The abundance of *Rikenellaceae_RC9_gut_group* is known to differ based on diet composition and has been observed to have a negative correlation with average daily gain (ADG) (Mcloughlin et al., 2020). However, current literature does not describe clear association between *Rikenellaceae_RC9_gut_group* and water intake or feed intake. Hence, the statistically significant decrease we observed during our study breaks new ground on this regard. However, we did not observe a significant increase in the abundance of *Rikenellaceae_RC9_gut_group* in animals with low percentage recovery of ADG during water restriction, as one might infer based on the previous observations of a negative correlation between *Rikenellaceae_RC9_gut_group* abundance and ADG.

2.4.4 The rumen core-microbiome remains essentially unchanged during water restriction.

In host-microbiome studies, a core microbiome is often defined to identify microbial taxa that plays a key role across time, treatment, spatial distribution and ecological impact etc. and usually consists of a persisting set of microbial taxa. Due to the lack of clear definitions, a common consensus on parameters and their threshold values to define the scope of a core-microbiome has not been established yet (Risely, 2020). Prevalence and relative abundance are two of the most common parameters used in host-microbiome studies to define a host-associated core-microbiome, even though the threshold levels may vary (John Wallace et al., 2019).

We used 70% prevalence and 0.01% relative abundance as threshold values to define the baseline and restriction rumen core-microbiomes. Meaning that a taxon will only be considered as a part of the core microbiome if it is present in at least 70% baseline/restriction rumen samples at an individual relative abundance > 0.01%.

The core microbiome derived for baseline rumen samples comprised of 9 genera, and the core microbiome for the restriction rumen samples consisted of 10 genera essentially identical to those of the baseline core microbiome, with the exception of the addition of *Lachnospiraceae_XPB1014_group. Ruminococcus*, *Prevotella*, and Lachnospiraceae family have previously been found to be a part of the heritable core microbiome of cattle (John Wallace et al., 2019). In a temporal study conducted using Tibetan lambs, *Christensenellaceae_R_7_group, Rikenellaceae_RC9_gut_group,* and

Prevotellaceae_UCG_003 were observed to get established in the rumen microbiome of mature lambs (Wang et al., 2019). To date, *Christensenellaceae minuta* which is closely related to the *Christensenellaceae_R_7_group*, is the only cultivated member of the *Christensenella* genus, known to have the capability of producing acetate and butyrate from glucose. Members of the *Christensenellacea* family are also known to be associated with lowering the pH of the rumen (Holman and Gzyl, 2019).

Previous work has found bacterial taxa such as *Prevotella, Butyrivibrio*, and *Ruminococcus*, and unclassified Lachnospiraceae as prominent members of the rumen microbiome, hence deemed as members of the rumen core microbiome at genus (or higher) level (Henderson et al. 2015). However, it is important to note that in our study, *Butyrivibrio* was not present in high abundance and prevalence in either baseline or restriction rumen samples, hence not being identified as a part of the core rumen microbiome, and not appearing in the 10 most abundant genera found in the rumen samples. It ranked 45th in overall abundance and had a baseline relative abundance of 0.29% and a restriction relative abundance of 0.24%.

2.4.5 LEfSe analysis and Random Forest classification reveal biomarkers associated with rumen microbiome under water stress.

Linear discriminant analysis (LDA) effect size or LEfSe (Segata et al., 2011) is a biomarker discovery algorithm specifically designed to identify biomarkers in high dimensional data such as metagenomics or microbiome data. Conventionally, the statistical significance of $p<0.05$ and a minimum LDA score of 2.0 are used as threshold

values for a feature to be identified as a biomarker. To identify biomarkers for baseline and restriction rumen samples, we used an p-value <0.05 and LDA>2.0 with all LEfSe analyses done at genus level (Chong et al., 2020).

Based on both LEfSe analysis and RF classification, a consensus list of 5 differentially abundant biomarkers were associated with restriction rumen samples, namely; *UCG_004, Prevotellaceae_UCG_004, Moryella, Lachnobacterium,* and *Papillibacter*. All of these biomarkers exhibited statistically significant ($p<0.05$) increase in relative abundances in restriction rumen samples when compared with baseline samples. *Moryella* is known to be positively associated with dry matter intake (DMI) (Qiu et al., 2019). The abundance of *Lachnobacterium* and *Papillibacter* (known to play a role in butyrate production in the rumen) are affected by diet type (Wang et al. 2020; Belanche et al. 2019). The abundance of *Prevotellaceae UCG-004* on the other hand is known to decrease during starvation. However, none of these taxa have been associated with water intake/water restriction so far. Hence the current study adds a new dimension to the existing knowledge we have on external factors that affect the abundance of rumen microbiota.

2.5 References

Ahlberg, C. M. 2018. Characterization of Water Intake in Beef Cattle: Test Length Guidelines, Water Intake Prediction, and Genetic Parameters. PhD Diss. Kansas State University, Manhattan.

Anderson, M. J. 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26:32–46. doi:10.1046/j.1442-9993.2001.01070.x.

Auffret, M. D., R. D. Stewart, R. J. Dewhurst, C. A. Duthie, M. Watson, and R. Roehe. 2020. Identification of Microbial Genetic Capacities and Potential Mechanisms Within the Rumen Microbiome Explaining Differences in Beef Cattle Feed Efficiency. Front. Microbiol. 11:1–16. doi:10.3389/fmicb.2020.01229.

Baek, Y. C., H. Choi, J.-Y. Jeong, S. D. Lee, M. J. Kim, S. Lee, S.-Y. Ji, and M. Kim. 2020. The impact of short-term acute heat stress on the rumen microbiome of Hanwoo steers. J. Anim. Sci. Technol. 62:208–217. doi:10.5187/jast.2020.62.2.208.

Belanche, A., A. H. Kingston-Smith, G. W. Griffith, and C. J. Newbold. 2019. A Multi-Kingdom Study Reveals the Plasticity of the Rumen Microbiota in Response to a Shift From Non-grazing to Grazing Diets in Sheep. Front. Microbiol. 10:122. doi:10.3389/fmicb.2019.00122.

Benatallah, A., F. Ghozlane, and M. Marie. 2019. The effect of water restriction on physiological and blood parameters in lactating dairy cows reared under Mediterranean climate. Asian-Australasian J. Anim. Sci. 32:152–158. doi:10.5713/ajas.18.0098.

Breiman, L. 2001. Random Forests. Mach. Learn. 45:5–32. doi:10.1023/A:1010933404324.

Bruno, K. A. 2019. Effects of Extended Water Restriction On Performance, Behavior, Health, And Welfare of Feedlot Steers. PhD Diss. Oklahoma State University, Stillwater.

Burgos, M. S., W. Langhans, and M. Senn. 2000. Role of rumen fluid hypertonicity in the dehydration-induced hypophagia of cows. Physiol. Behav. 71:423–430. doi:10.1016/S0031-9384(00)00357-7.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583. doi:10.1038/nmeth.3869.

Carberry, C. A., S. M. Waters, D. A. Kenny, and C. J. Creevey. 2014. Rumen methanogenic genotypes differ in abundance according to host residual feed intake phenotype and diet type. Appl. Environ. Microbiol. 80:586–594. doi:10.1128/AEM.03131-13.

Carter, R. R., and W. L. Grovum. 1990. A review of the physiological significance of hypertonic body fluids on feed intake and ruminal function: salivation, motility and microbes. J. Anim. Sci. 68:2811–2832. doi:10.2527/1990.6892811x.

Chong, J., P. Liu, G. Zhou, and J. Xia. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat. Protoc. 15:799–821. doi:10.1038/s41596-019-0264-1.

Clemmons, B. A., S. T. Reese, F. G. Dantas, G. A. Franco, T. P. L. Smith, O. I. Adeyosoye, K. G. Pohler, and P. R. Myer. 2017. Vaginal and uterine bacterial communities in postpartum lactating cows. Front. Microbiol. 8:1–10. doi:10.3389/fmicb.2017.01047.

Delgado, B., A. Bach, I. Guasch, C. González, G. Elcoso, J. E. Pryce, and O. Gonzalez-Recio. 2019. Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake levels in cattle. Sci. Rep. 9:1–13. doi:10.1038/s41598-018-36673-w.

Dethlefsen, L., S. Huse, M. L. Sogin, and D. A. Relman. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16s rRNA sequencing. PLoS Biol. 6:2383–2400. doi:10.1371/journal.pbio.0060280.

Dhariwal, A., J. Chong, S. Habib, I. L. King, L. B. Agellon, and J. Xia. 2017. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 45:W180–W188. doi:10.1093/nar/gkx295.

Dyer, T., J. Rossy, and M. Pence. 2017. Water Requirements and Quality Issues for Cattle. Available from: https://secure.caes.uga.edu/extension/publications/files/pdf/SB56_5.PDF

Fan, P., B. Bian, L. Teng, C. D. Nelson, J. Driver, M. A. Elzo, and K. C. Jeong. 2020. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 14:302–317. doi:10.1038/s41396-019-0529-2.

Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: And this is not optional. Front. Microbiol. 8:1–6. doi:10.3389/fmicb.2017.02224.

Hackmann, T. J., and J. L. Firkins. 2015. Maximizing efficiency of rumen microbial protein production. Front. Microbiol. 6:1–16. doi:10.3389/fmicb.2015.00465.

Henderson, G., F. Cox, S. Ganesh, A. Jonker, W. Young, P. H. Janssen, L. Abecia, E. Angarita, P. Aravena, G. N. Arenas, C. Ariza, G. T. Attwood, J. M. Avila, J. Avila-Stagno, A. Bannink, R. Barahona, M. Batistotti, M. F. Bertelsen, A. Brown-Kav, A. M. Carvajal, L. Cersosimo, A. V. Chaves, J. Church, N. Clipson, M. A. Cobos-Peralta, A. L.

Cookson, S. Cravero, O. C. Carballo, K. Crosley, G. Cruz, M. C. Cucchi, R. De La Barra, A. B. De Menezes, E. Detmann, K. Dieho, J. Dijkstra, W. L. S. Dos Reis, M. E. R. Dugan, S. H. Ebrahimi, E. Eythórsdóttir, F. N. Fon, M. Fraga, F. Franco, C. Friedeman, N. Fukuma, D. Gagić, I. Gangnat, D. J. Grilli, L. L. Guan, V. H. Miri, E. Hernandez-Sanabria, A. X. I. Gomez, O. A. Isah, S. Ishaq, E. Jami, J. Jelincic, J. Kantanen, W. J. Kelly, S. H. Kim, A. Klieve, Y. Kobayashi, S. Koike, J. Kopecny, T. N. Kristensen, S. J. Krizsan, H. LaChance, M. Lachman, W. R. Lamberson, S. Lambie, J. Lassen, S. C. Leahy, S. S. Lee, F. Leiber, E. Lewis, B. Lin, R. Lira, P. Lund, E. Macipe, L. L. Mamuad, H. C. Mantovani, G. A. Marcoppido, C. Márquez, C. Martin, G. Martinez, M. E. Martinez, O. L. Mayorga, T. A. McAllister, C. McSweeney, L. Mestre, E. Minnee, M. Mitsumori, I. Mizrahi, I. Molina, A. Muenger, C. Munoz, B. Murovec, J. Newbold, V. Nsereko, et al. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5. doi:10.1038/srep14567.

Hill, J., C. McSweeney, A. D. G. Wright, G. Bishop-Hurley, and K. Kalantar-zadeh. 2016. Measuring Methane Production from Ruminants. Trends in Biotechnol. 34:26–35. doi:10.1016/j.tibtech.2015.10.004.

Holman, D. B., and K. E. Gzyl. 2019. A meta-analysis of the bovine gastrointestinal tract microbiota. FEMS Microbiol. Ecol. 95:1–9. doi:10.1093/femsec/fiz072.

Huws, S. A., C. J. Creevey, L. B. Oyama, I. Mizrahi, S. E. Denman, M. Popova, R. Muñoz-Tamayo, E. Forano, S. M. Waters, M. Hess, I. Tapio, H. Smidt, S. J. Krizsan, D. R. Yáñez-Ruiz, A. Belanche, L. Guan, R. J. Gruninger, T. A. McAllister, C. J. Newbold, R. Roehe, R. J. Dewhurst, T. J. Snelling, M. Watson, G. Suen, E. H. Hart, A. H. Kingston-Smith, N. D. Scollan, R. M. Do Prado, E. J. Pilau, H. C. Mantovani, G. T. Attwood, J. E. Edwards, N. R. McEwan, S. Morrisson, O. L. Mayorga, C. Elliott, and D. P. Morgavi. 2018. Addressing global ruminant agricultural challenges through understanding the rumen microbiome: Past, present, and future. Front. Microbiol. 9:1–33. doi:10.3389/fmicb.2018.02161.

John Wallace, R., G. Sasson, P. C. Garnsworthy, I. Tapio, E. Gregson, P. Bani, P. Huhtanen, A. R. Bayat, F. Strozzi, F. Biscarini, T. J. Snelling, N. Saunders, S. L. Potterton, J. Craigon, A. Minuti, E. Trevisi, M. L. Callegari, F. P. Cappelli, E. H. Cabezas-Garcia, J. Vilkki, C. Pinares-Patino, K. O. Fliegerová, J. Mrázek, H. Sechovcová, J. Kopečný, A. Bonin, F. Boyer, P. Taberlet, F. Kokou, E. Halperin, J. L. Williams, K. J. Shingfield, and I. Mizrahi. 2019. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. Sci. Adv. 5. doi:10.1126/sciadv.aav8391.

Kandlikar, G. S., Z. J. Gold, M. C. Cowen, R. S. Meyer, A. C. Freise, N. J. B. Kraft, J. Moberg-Parker, J. Sprague, D. J. Kushner, and E. E. Curd. 2018. Ranacapa: An R package and shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. F1000Research. 7:1–18. doi:10.12688/f1000research.16680.1.

Klevenhusen, F., R. M. Petri, M. T. Kleefisch, R. Khiaosa-ard, B. U. Metzler-Zebeli, and Q. Zebeli. 2017. Changes in fibre-adherent and fluid-associated microbial communities

and fermentation profiles in the rumen of cattle fed diets differing in hay quality and concentrate amount. FEMS Microbiol. Ecol. 93. doi:10.1093/femsec/fix100.

Li, F., C. Li, Y. Chen, J. Liu, C. Zhang, B. Irving, C. Fitzsimmons, G. Plastow, and L. L. Guan. 2019. Host genetics influence the rumen microbiota and heritable rumen microbial features associate with feed efficiency in cattle. Microbiome. 7:1–17. doi:10.1186/s40168-019-0699-1.

Lim, M. Y., E. J. Song, S. H. Kim, J. Lee, and Y. Do Nam. 2018. Comparison of DNA extraction methods for human gut microbial community profiling. Syst. Appl. Microbiol. 41:151–157. doi:10.1016/j.syapm.2017.11.008.

Liu, C., Q. Meng, Y. Chen, M. Xu, M. Shen, R. Gao, and S. Gan. 2017. Role of age-related shifts in rumen bacteria and methanogens in methane production in cattle. Front. Microbiol. 8:1–14. doi:10.3389/fmicb.2017.01563.

Luz Calle, M. 2019. Statistical analysis of metagenomics data. Genomics and Informatics. 17. doi:10.5808/GI.2019.17.1.e6.

Matthews, C., F. Crispie, E. Lewis, M. Reid, P. W. O'Toole, and P. D. Cotter. 2019. The rumen microbiome: a crucial consideration when optimising milk and meat production and nitrogen utilisation efficiency. Gut Microbes. 10:115–132. doi:10.1080/19490976.2018.1505176.

Mcloughlin, S., C. Spillane, N. Claffey, P. E. Smith, T. O. Rourke, M. G. Diskin, and S. M. Waters. 2020. Rumen Microbiome Composition Is Altered in Sheep Divergent in Feed Efficiency. Front. Microbiol. 11:1–16. doi:10.3389/fmicb.2020.01981.

McMurdie, P. J., and S. Holmes. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One. 8. doi:10.1371/journal.pone.0061217.

McMurdie, P. J., and S. Holmes. 2014. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Comput. Biol. 10. doi:10.1371/journal.pcbi.1003531.

Moore, S. G., A. C. Ericsson, S. E. Poock, P. Melendez, and M. C. Lucy. 2017. Hot topic: 16S rRNA gene sequencing reveals the microbiome of the virgin and pregnant bovine uterus. J. Dairy Sci. 100:4953–4960. doi:10.3168/jds.2017-12592.

Nardone, A., B. Ronchi, N. Lacetera, M. S. Ranieri, and U. Bernabucci. 2010. Effects of climate changes on animal production and sustainability of livestock systems. Livest. Sci. 130:57–69. doi:10.1016/j.livsci.2010.02.011.

Ocejo, M., B. Oporto, and A. Hurtado. 2019. 16S rRNA amplicon sequencing characterization of caecal microbiome composition of broilers and free-range slow-growing chickens throughout their productive lifespan. Sci. Rep. 9:2506. doi:10.1038/s41598-019-39323-x.

Parsons, B. N., U. Z. Ijaz, R. D'Amore, M. D. Burkitt, R. Eccles, L. Lenzi, C. A. Duckworth, A. R. Moore, L. Tiszlavicz, A. Varro, N. Hall, and D. M. Pritchard. 2017.

Comparison of the human gastric microbiota in hypochlorhydric states arising as a result of Helicobacter pylori-induced atrophic gastritis, autoimmune atrophic gastritis and proton pump inhibitor use. bioRxiv. 1–18. doi:10.1101/144907.

Pitta, D. W., W. E. Pinchak, S. E. Dowd, J. Osterstock, V. Gontcharova, E. Youn, K. Dorton, I. Yoon, B. R. Min, J. D. Fulford, T. A. Wickersham, and D. P. Malinowski. 2010. Rumen bacterial diversity dynamics associated with changing from bermudagrass hay to grazed winter wheat diets. Microb. Ecol. 59:511–522. doi:10.1007/s00248-009-9609-6.

Pitta, D. W., W. E. Pinchak, N. Indugu, B. Vecchiarelli, R. Sinha, and J. D. Fulford. 2016. Metagenomic Analysis of the Rumen Microbiome of Steers with Wheat-Induced Frothy Bloat. Front. Microbiol. 7:689. doi:10.3389/fmicb.2016.00689.

Qiu, Q., C. Gao, Z. Gao, M. A. ur Rahman, Y. He, B. Cao, and H. Su. 2019. Temporal Dynamics in Rumen Bacterial Community Composition of Finishing Steers during an Adaptation Period of Three Months. Microorganisms. 7:410. doi:10.3390/microorganisms7100410.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41:590–596. doi:10.1093/nar/gks1219.

Rabaza, A., G. Banchero, C. Cajarville, P. Zunino, A. Britos, J. L. Repetto, and M. Fraga. 2019. Effects of feed withdrawal duration on animal behaviour, rumen microbiota and blood chemistry in feedlot cattle: Implications for rumen acidosis. Animal. 14:66–77. doi:10.1017/S1751731119001538.

Ramette, A. 2007. Multivariate analyses in microbial ecology. FEMS Microbiol. Ecol. 62:142–160. doi:10.1111/j.1574-6941.2007.00375.x.

Rasby, R. J., and T. M. Walz. 2011. Water Requirements for Beef Cattle. Available from: http://extension.unl.edu/publications.

Risely, A. 2020. Applying the core microbiome to understand host–microbe systems. J. Anim. Ecol. 89:1549–1558. doi:10.1111/1365-2656.13229.

Roehe, R., R. J. Dewhurst, C. A. Duthie, J. A. Rooke, N. McKain, D. W. Ross, J. J. Hyslop, A. Waterhouse, T. C. Freeman, M. Watson, and R. J. Wallace. 2016. Bovine Host Genetic Variation Influences Rumen Microbial Methane Production with Best Selection Criterion for Low Methane Emitting and Efficiently Feed Converting Hosts Based on Metagenomic Gene Abundance. PLoS Genet. 12:1–20. doi:10.1371/journal.pgen.1005846.

Roguet, A., A. M. Eren, R. J. Newton, and S. L. McLellan. 2018. Fecal source identification using random forest. Microbiome. 6:1–15. doi:10.1186/s40168-018-0568-3.

Romero-Pérez, G. A., K. H. Ominski, T. A. McAllister, and D. O. Krause. 2011. Effect of environmental factors and influence of rumen and hindgut biogeography on bacterial

communities in steers. Appl. Environ. Microbiol. 77:258–268. doi:10.1128/AEM.01289-09.

Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, and C. Huttenhower. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60. doi:10.1186/gb-2011-12-6-r60.

Seshadri, R., S. C. Leahy, G. T. Attwood, K. H. Teh, S. C. Lambie, A. L. Cookson, E. A. Eloe-Fadrosh, G. A. Pavlopoulos, M. Hadjithomas, N. J. Varghese, D. Paez-Espino, R. Perry, G. Henderson, C. J. Creevey, N. Terrapon, P. Lapebie, E. Drula, V. Lombard, E. Rubin, N. C. Kyrpides, B. Henrissat, T. Woyke, N. N. Ivanova, W. J. Kelly, N. Palevic, P. H. Janssen, R. S. Ronimus, S. Noel, P. Soni, K. Reilly, T. Atherly, C. Ziemer, A. D. Wright, S. Ishaq, M. Cotta, S. Thompson, K. Crosley, N. McKain, J. J. Wallace, H. J. Flint, J. C. Martin, R. J. Forster, R. J. Gruninger, T. McAllister, R. Gilbert, D. J. Ouwerkerk, A. J. Klieve, R. Al Jassim, S. Denman, C. McSweeney, C. Rosewarne, S. Koike, Y. Kobayashi, M. Mitsumori, T. Shinkai, S. Cravero, and M. Cerón Cucchi. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. Nat. Biotechnol. 36:359–367. doi:10.1038/nbt.4110.

Smith, B. 2011. Optimising Bacterial DNA Extraction from Faecal Samples: Comparison of Three Methods. Open Microbiol. J. 5:14–17. doi:10.2174/1874285801105010014.

Snelling, T. J., M. D. Auffret, C.-A. Duthie, R. D. Stewart, M. Watson, R. J. Dewhurst, R. Roehe, and A. W. Walker. 2019. Temporal stability of the rumen microbiota in beef cattle, and response to diet and supplements. Anim. Microbiome. 1:1–14. doi:10.1186/s42523-019-0018-y.

Stein, R. A., and D. E. Katz. 2017. Escherichia coli, cattle and the propagation of disease. FEMS Microbiol. Lett. 364:50. doi:10.1093/femsle/fnx050.

Tajima, K., I. Nonaka, K. Higuchi, N. Takusari, M. Kurihara, A. Takenaka, M. Mitsumori, H. Kajikawa, and R. I. Aminov. 2007. Influence of high temperature and humidity on rumen bacterial diversity in Holstein heifers. Anaerobe. 13:57–64. doi:10.1016/j.anaerobe.2006.12.001.

Wagner, B. D., G. K. Grunwald, G. O. Zerbe, S. K. Mikulich-Gilbertson, C. E. Robertson, E. T. Zemanick, and J. K. Harris. 2018. On the use of diversity measures in longitudinal sequencing studies of microbial communities. Front. Microbiol. 9. doi:10.3389/fmicb.2018.01037.

Wang, H., F. Wu, T. Guan, Y. Zhu, Z. Yu, D. Zhang, S. Zhang, H. Su, and B. Cao. 2020. Chopping Roughage Length Improved Rumen Development of Weaned Calves as Revealed by Rumen Fermentation and Bacterial Community. Animals. 10:2149. doi:10.3390/ani10112149.

Wang, L., K. Zhang, C. Zhang, Y. Feng, and X. Zhang. 2019. Dynamics and stabilization of the rumen microbiome in yearling Tibetan sheep. Sci. Rep. 1–9. doi:10.1038/s41598-019-56206-3.

Weimer, P. J., J. B. Russell, and R. E. Muck. 2009. Lessons from the cow: What the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass.

Bioresour. Technol. 100:5323–5331. doi:10.1016/j.biortech.2009.04.075.

Wickramasinghe, H. K. J. P., A. J. Kramer, and J. A. D. R. N. Appuhamy. 2019. Drinking water intake of newborn dairy calves and its effects on feed intake, growth performance, health status, and nutrient digestibility. J. Dairy Sci. 102:377–387. doi:10.3168/jds.2018-15579.

Willis, A. D. 2019. Rarefaction, alpha diversity, and statistics. Front. Microbiol. 10. doi:10.3389/fmicb.2019.02407.

Zhao, S., L. Min, N. Zheng, and J. Wang. 2019. Effect of heat stress on bacterial composition and metabolism in the rumen of lactating dairy cows. Animals. 9. doi:10.3390/ani9110925.

CHAPTER III

POPULATION DYNAMICS OF THE BOVINE FECAL MICROBIOME UNDER WATER RESTRICTED CONDITIONS: A TAXONOMIC SURVEY OF 16S V4 AMPLICON SEQUECING DATA

3.1 Introduction

The cattle gut microbiome and its host-microbiome interactions has been studied extensively in order to develop new strategies to improve production and efficiency of beef and dairy cattle systems while reducing its negative environmental impact (O'Hara et al., 2020). The rumen microbiome has often been the epicenter of cattle gut microbiome studies due to its importance in fermentation and VFA production. With the increased use of Next Generation Sequencing (NGS) technologies, such studies have generated a vast amount of data and information detailing taxonomic composition and function of the rumen microbiome, spanning multiple breeds, production systems, and external environmental factors (Henderson et al., 2015; Deusch et al., 2017; Zhou et al., 2018).

The lower gut of cattle comprises of the small intestine and the hindgut region comprising of cecum, colon and the rectum. The hindgut is thought to be responsible for up to 30% of the total cellulose and hemicellulose digestion in ruminants (Hoover 1978; Gressley et al. 2011), and the energy produced can be an important contributor to the total energy requirement of the animal throughout different stages of production,

especially during the first few weeks of the life of a calf where the rumen is not yet fully developed (Castro et al., 2016). When compared to the rumen microbiome, the structure and function of the lower-gut microbiome and its role in animal health and production is not well understood (O'Hara et al., 2020). However, recent studies suggest that the inherent microbiota of the lower gastrointestinal (GI) tract may play an important role in beef and dairy cattle production, including host-microbiome interactions that contribute to animal health, even though comprehensive studies on this regard is still scarce (Malmuthuge and Guan 2017; Myer et al. 2015; Hara et al. 2018). In contrast to the rumen, the lower gut of ruminants plays an imperative role in their immune function. The mucosal epithelium hosting a multitude of immune receptors and cells, acts as a chemical and physical barrier for pathogenic organisms (Hooper et al., 2015; Malmuthuge and Guan, 2017). Similar to their function in the monogastrics, the lower gut microbial community of cattle is also found to contribute towards the establishment of the immune system (Mulder et al., 2011; Malmuthuge and Guan, 2017). Therefore, gaining a comprehensive understanding of structure and function of the fecal microbiome of cattle is of utmost importance due its potential role in improving ruminant gut health, maximizing production efficiency and reducing the environmental impact of beef cattle operations (O'Hara et al., 2020).

Analyzing the fecal microbial community has often been used as a non-invasive method of studying the hindgut microbiome (Mote et al., 2019) especially due to the similarities in the composition of microbial communities recovered from cattle feces, and their hindgut digesta (Song et al., 2018). Existing studies have found the cattle fecal

microbiome to comprise of a complex microbial community, and like the rumen microbiome, is often affected by both host and external environmental factors such as breed, age, stage of production, and diet (Shanks et al., 2011; Song et al., 2018; Hagey et al., 2019; Noel et al., 2019; Cendron, 2020). However, the effect of water intake or water restriction on the establishment and development of the lower gut microbiome of cattle is yet to be extensively studied.

Studies conducted in mice and humans have demonstrated that pH and chlorine concentration of drinking water can have a significant impact in shaping the gut microbiome (Sofi et al., 2014; Sasada et al., 2015; Cremer et al., 2017), thus establishing the fact that water intake can have a significant impact on shaping the mammalian gut microbiome. Therefore, a prominent research gap exists in studying the effects of water restriction or water stress on the fecal microbiome of cattle thus justifying a comprehensive microbiome study on this regard.

Hence, we conducted the current study to elucidate the microbial dynamics of the bovine lower gut microbiome under water stress, based on 16S rRNA gene amplicon sequencing of genomic DNA extracted from fecal samples obtained from feedlot beef cattle under *ad libitum* water intake, and at 50% water restriction.

3.2 Materials and Methods

3.2.1 Animals, experimental design and samples collection

The animals and the experimental design used in this study were as described in Chapter 2. The difference lies in sample collection where we collected fecal samples instead of rumen samples, in order to investigate the fecal microbiome. Refer section 2.2.1 and 2.2.2 for detailed information. Fecal samples were collected directly from the rectum (being careful not to cross contaminate samples), at the end of each trial. All samples were snap frozen on dry ice at sample collection, transported immediately to the Lab and stored at -80C^0 until DNA extraction.

3.2.2 Animal selection for DNA extraction

Even though fecal sample collection was attempted on all animals, DNA extraction (for 16S rRNA gene amplicon sequencing) from all fecal samples is not cost effective. Hence, animals that showed statistically significant ($p < 0.1$) performance based on percentage recovery of average daily gain (ADG) were selected, and their fecal samples were used for genomic DNA extraction.

$$\% \text{ recovery of ADG} = \frac{\text{ADG during restriction trial}}{\text{ADG during WI trial}} \times 100$$

The % recovery of ADG distributions were calculated separately for each group and the animals that fall in the tails of the distribution were selected to maximize statistical power. A total of 146 animals from the seven groups were selected for fecal DNA extraction based on the criteria mentioned above. A total of 223 fecal samples that

comprised of 118 samples from the water intake trial (baseline fecal samples) and 115 samples from the water restriction trial (restriction fecal samples) originating from the 146 animals were used for the current study.

3.2.3 Whole genomic DNA extraction from fecal samples

Based on the results from the pilot study, QIAamp® PowerFecal Pro DNA Kit with standard Power Beads was used to extract genomic DNA from fecal samples according to manufacturer's specifications subjected to modifications in the bead beating and final elution steps. Bead beating was done using the BeadBug6® homogenizer (Benchmark Scientific, Inc. Edison, NJ) for two 2min iterations at an rpm of 4000 separated by a resting period of 5mins on ice (Smith, 2011; Lim et al., 2018). The BeadBug program was set to 4 cycles of 30sec each. With a 30sec resting period.

Final elution was done using molecular grade water as two iterations of 100µl, each with a 5min incubation at room temperature. DNA was eluted by centrifuging at 14000 rpm for 1min. The second elution was concentrated using the CentriVap DNA concentrator (Labconco, Kansas City, MO) by evaporating it down to 50µl. The second elution was then added to the first elution to arrive at a volume of 150µl in the final DNA elution.

3.2.4 Library preparation and 16S V4 rRNA gene amplicon sequencing

The PCR amplification of extracted DNA and 16S amplicon sequencing was performed by Novogene Corporation Inc, Sacramento, CA. The V4 region of the 16S ribosomal RNA gene (~250bp) was amplified using the specific primer pair 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGAC TACHVGGGTWTCTAAT-

3'), with the forward primer carrying the barcode sequence. Phusion® High-Fidelity PCR Master Mix (New England Biolabs) was used to carry out all PCR reactions.

The PCR products were purified by running them on a 2% agarose gel and extracting them using a Qiagen® Gel Extraction Kit (Qiagen, Germany). The sequencing library preparation was done using a NEBNext® UltraTM DNA Library Prep Kit for Illumina. It was quality tested using Qubit 3.0 fluorometer and quantified using Q-PCR. Subsequently, an Illumina HiSeq 2500 platform was utilized for the sequencing of the library, generating ~250bp paired-end raw sequence reads.

3.2.5 16S V4 amplicon sequencing data analysis

The computational analysis of 16S V4 amplicon sequence data was conducted using DADA2 version 1.8, according to standard operation procedures for Illumina data (Callahan et al., 2016). All DADA2 codes were run using R version 3.6.1. on the Oklahoma State University high performance computing cluster (Pete). Briefly, the quality scores of forward and reverse sequence reads were plotted to identify any drastic drops in sequence quality, specifically towards the ends of the sequences, and no such significant drops were observed. Hence, no specific truncation of the ends of sequences were done during the pre-processing step based on sequence quality. However, linker primer sequences and any sequences that had ambiguous bases and an expected error greater than 2 were removed using the 'filterAndTrim' command in DADA2. Subsequently, the error rates for the amplicon sequence data were modeled using 'learnErrors' function in DADA2 and the sequence reads were then de-replicated using the function 'derepFastq'. The de-replicated sequences were then used for the inference

of sequence variants for each sample using the core sample inference algorithms available in DADA2, implemented in the function ‘dada’. Using the ‘mergePairs’ function, the corresponding forward and reverse inferred sequence variants were then merged to arrive at the full de-noised sequence contigs. The ‘makeSequenceTable’ function was run on these contigs to create the amplicon sequence variant (ASV) table. The ASV table was then subjected to chimera removal and taxonomic classification via ‘removeBimeraDenovo’ and ‘assignTaxonomy’ functions. The DADA formatted version of Silva reference database version 132 (Quast et al., 2013) was used as the training dataset for taxonomic classification. The ASV table and the taxonomy file generated by DADA2 was then used for statistical analysis and data visualization, using MicrobiomeAnalyst (Dhariwal et al., 2017).

3.2.6 Data filtering, normalization, statistical analysis and visualization

Low abundance features, defined as those with less than 10 read counts in at least 20% of the samples were removed prior to data normalization. Data rarefying (based on minimum library size), and Total Sum Scaling was performed as data normalization steps to negate the effect of uneven sequencing depth (Dhariwal et al., 2017). The resulting filtered and normalized data were used for subsequent statistical analyses and data visualizations. Baseline fecal samples from all selected animals were grouped together when making inferences about the fecal microbiome during *ad libitum* water intake and restriction fecal samples from all animals were grouped together when making inferences about the fecal microbiome during water restriction.

3.3 Results

3.3.1 Rarefaction analysis

A total of 9,797,366 high quality sequences (86,702 ± 24,738) ranging from a minimum of 30,357 in a sample to a maximum of 123,556, were obtained from fecal restriction samples. For fecal baseline samples, a total of 10,134,524 high quality sequences were obtained with a minimum sequencing read depth of 27,215 to a maximum read depth of 124,731 (85,886 ± 22,822)

For the purpose of determining whether enough sequencing depth was achieved to capture the entirety of the microbial richness in baseline and restriction fecal samples during 16S rRNA gene amplicons sequencing, we conducted a rarefaction analysis to generate rarefaction curves and Good's coverage (Chao and Jost, 2012) values for each fecal sample. The rarefaction curves generated for samples from both water intake (baseline) and water restriction trials are depicted in Figure 3.1. Plateaued curves were observed for all fecal samples, well before the count of re-sampled sequences equaled minimum library size (27,261 read counts). For all fecal samples, a Good's coverage of >99% was achieved when re-sampled read counts reached 25,634.

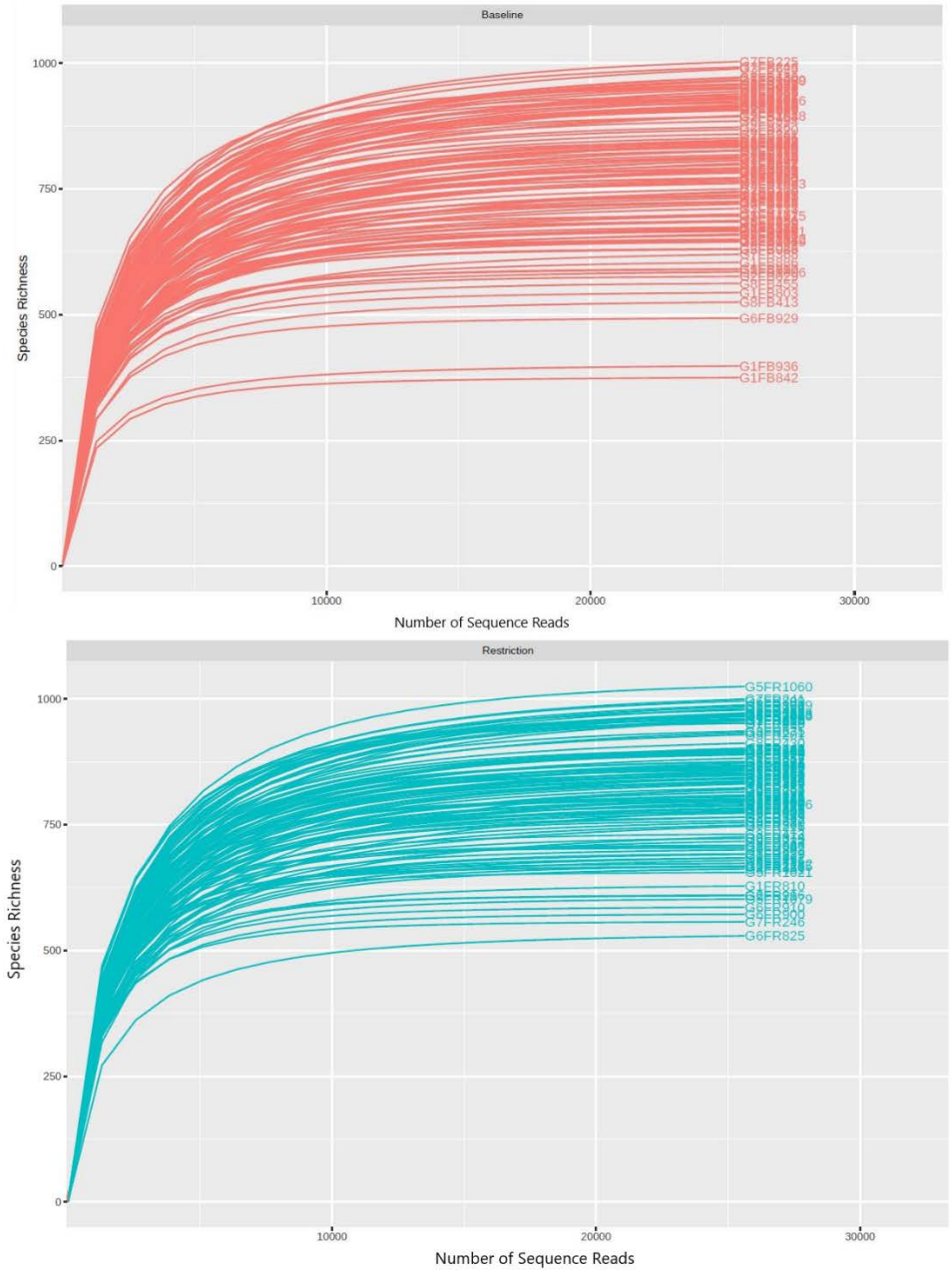


Figure 3.1: Rarefaction curves for baseline (top) and restriction (bottom) fecal samples. The curves were generated using the 'ggrare' function in ranacapa R pckage (Kandlikar et al., 2018) implemented in MicrobiomeAnalyst (Chong et al., 2020), and demonstrates how the discovery of new species (species richness in y axis) increases with the number of sampled reads from each sample (x axis). All curves are truncated at a sampling depth

of 27,261 sequence reads due to subsampling based on the minimum library size for all fecal samples (data rarefying).

3.3.2 A comparison of alpha diversity, between restriction and baseline fecal microbiomes.

As a part of microbial community profiling of fecal samples, and to determine whether water restriction has had significant impact on it, we conducted an alpha diversity analysis on both baseline and restriction fecal DNA samples. Chao1 and Observed ASV indices were used to calculate alpha diversity based on species richness (Figure 3.2). Shannon, Simpsons, and Fisher indices were used to calculate alpha diversity mainly based on species evenness (Ocejo et al., 2019). Between group comparisons for Chao1, Observed ASV and Fisher's indices were performed using t-test. Mann-Whitney U test was used for between group comparisons of Shannon and Simpson alpha diversity estimates (Clemmons et al., 2017). The phyloseq R package implemented in MicrobiomeAnalyst (McMurdie and Holmes, 2013; Chong et al., 2020) was used to conduct the alpha diversity analysis.

As depicted in Table 3.1, we did not observe any statistically significant differences ($p<0.05$) between baseline and restriction fecal samples for Chao1, Observed ASV and Fisher's indices. However, for both Shannon and Simpson's indices, we were able to observe statistically significant results ($p<0.05$) for between group comparisons of alpha diversity values.

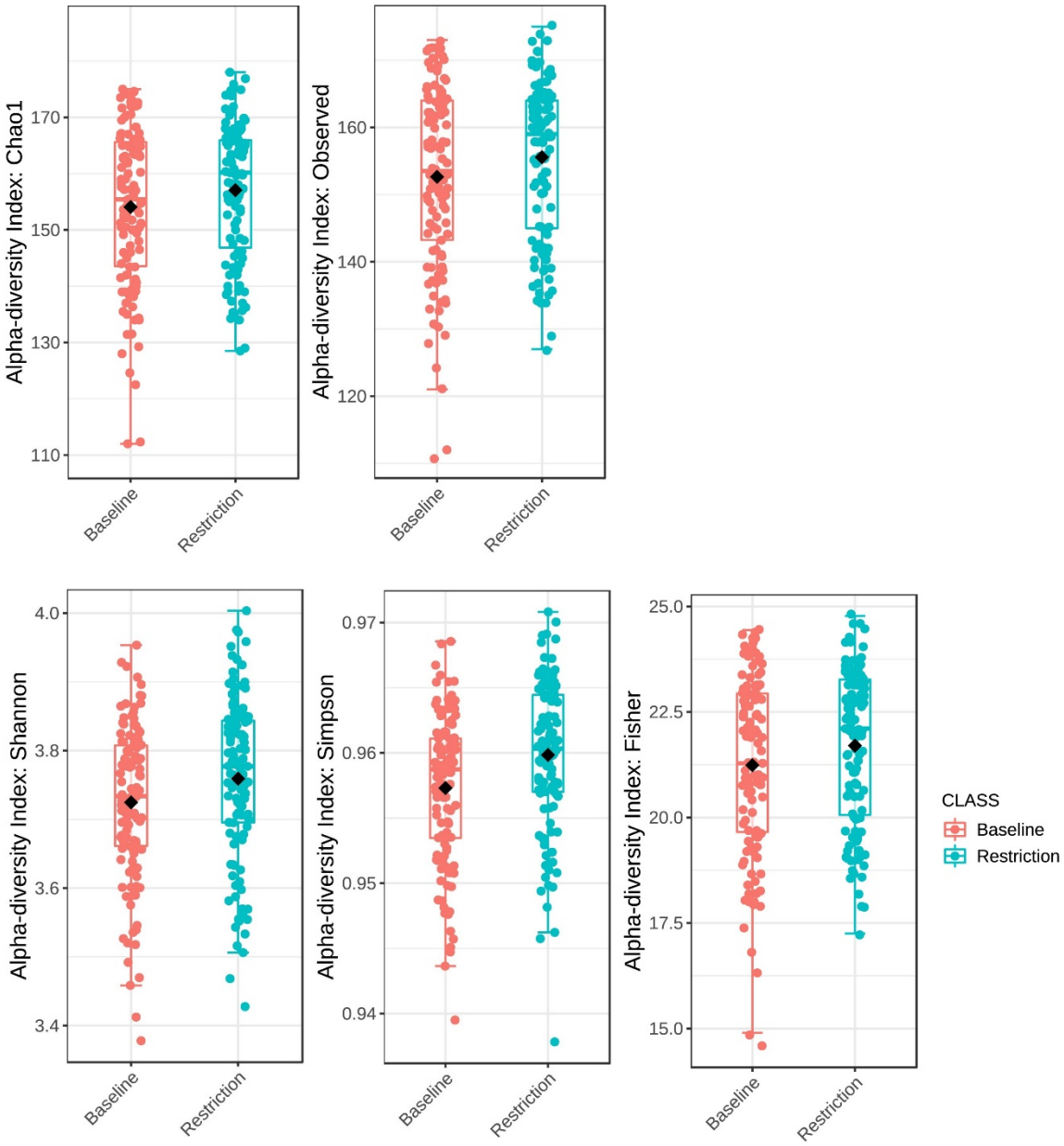


Figure 3.2: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from baseline and restriction fecal DNA samples. Each dot on the plot represents the alpha diversity of a single fecal sample. The dots are color coded based on which group they belong to (baseline or restriction). The y-axis indicates the magnitude of alpha diversity for each sample, based on the diversity measure used.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-1.7763	0.07701
Chao1	T-test	1.0429	0.06156
Shannon	Man-Whitney	5679	0.03162[a]
Simpson	Man-Whitney	5191	0.00195[b]
Fisher's	T-test	-1.8117	0.07136

(a) – $p<0.05$; (b) – $p<0.01$

Table 3.1: Between group comparisons for alpha diversity indices generated for fecal samples. T-test was used to measure between group comparisons for the distribution of Observed ASV, Chao1, and Fisher's indices. Man-Whitney test was used for Shannon and Simpson indices. Significance levels are indicated using 'a' and 'b' superscripts.

3.3.3 Baseline and restriction fecal microbiomes differ significantly in their beta diversity

We conducted a beta diversity analysis to assess the effect of water restriction on the species composition of the fecal microbial community, using abundance data from both baseline and restriction fecal samples. Two popularly used dissimilarity indices, namely; Bray-Curtis dissimilarity and Jensen-Shannon index (Chong et al., 2020) was used to measure dissimilarity/similarity between baseline and restriction microbial communities in each sample.

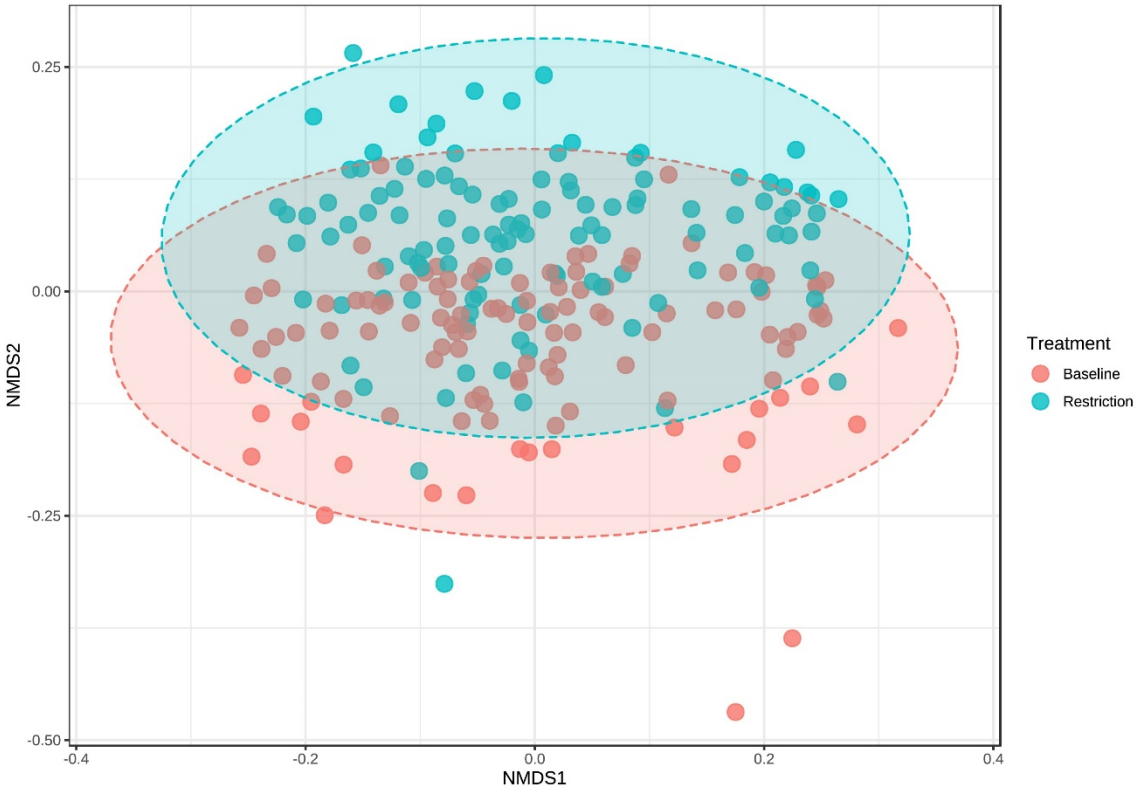
As illustrated in Figure 3.3 and 3.4, the beta diversity measures were then visualized using NMDS (Non-parametric multidimensional scaling) plots (Luz Calle, 2019). The baseline and restriction samples formed partially overlapping but distinguishable

clusters. We observed high statistical significance ($p<0.05$) when between group variance for the overall clustering pattern

was analyzed using permutational analysis of variance (PERMANOVA) (Moore et al., 2017). Statistics for between group comparisons are depicted in Table 2.2.

Distance method	**F-value**	**R-squared**	**P-value**	**NMDS Stress**
Bray-Curtis dissimilarity index	17.362	0.0699	$p < 0.001$	0.20326
Jensen-Shannon divergence	27.002	0.10466	$p < 0.001$	0.17904

Table 3.2: Results of the PERMANOVA analysis and the goodness-of-fit test for NMDS plot (NMDS Stress). The values are calculated using the underlying data generated using each of the distance methods listed.


0.25
0.00
-0.25
-0.50
NMDS2
-0.4
-0.2
0.0
0.2
0.4
NMDS1
Treatment
Baseline
Restriction

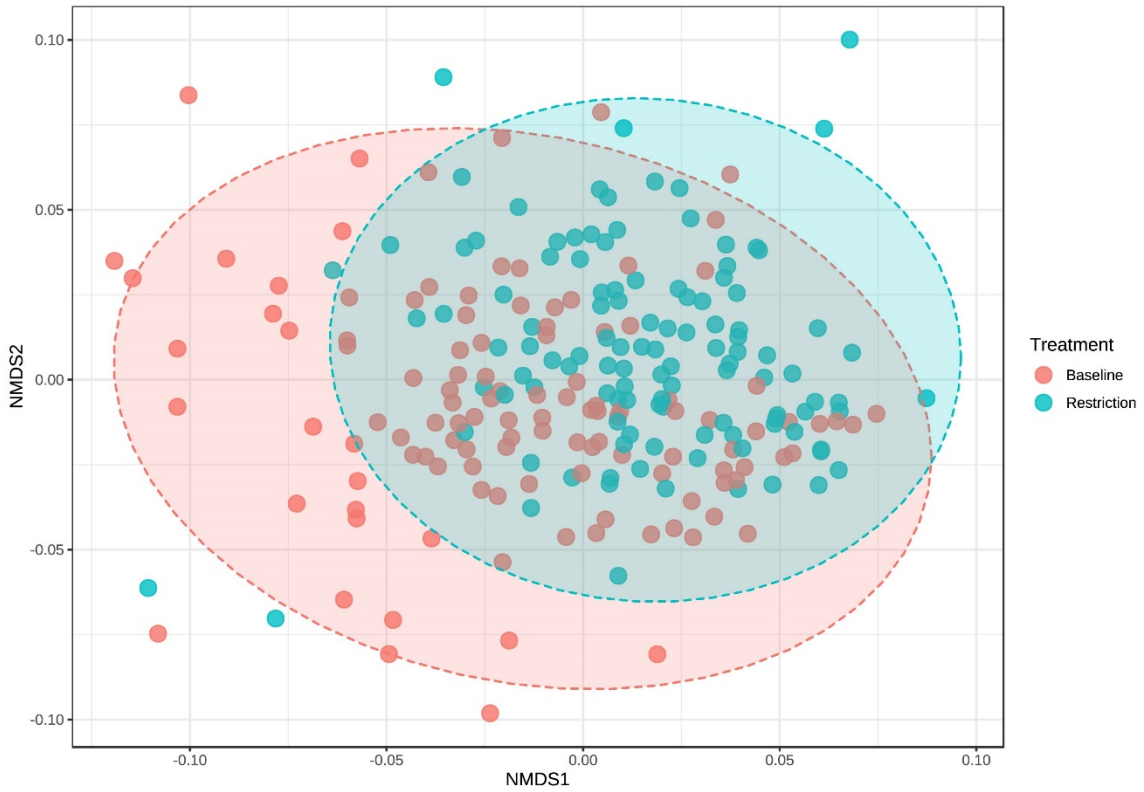

0.10
0.05
0.00
-0.05
-0.10
NMDS2
-0.10
-0.05
0.00
0.05
0.10
NMDS1
Treatment
Baseline
Restriction

Figure 3.3: NMDS plots illustrating the distribution of baseline and restriction fecal samples based on Bray-Curtis (a) and Jensen-Shannon (b) distance measures, in two-dimensional space. Each blue (restriction fecal samples) and red (baseline fecal samples) point represent the species composition of the entire microbiome present in a given fecal sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups.

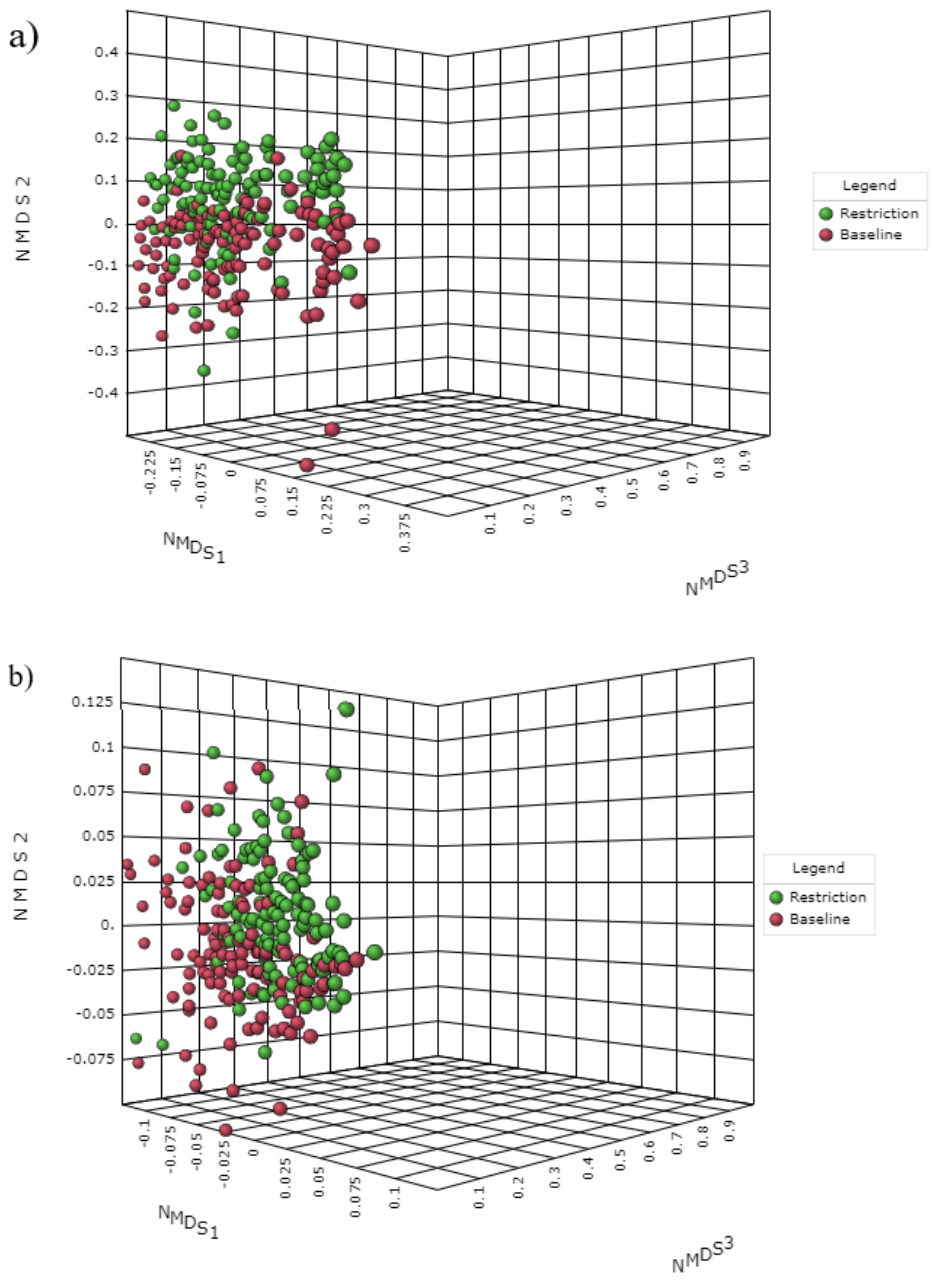


Figure 3.4: Three-dimensional NMDS plots illustrating the distribution of restriction and baseline fecal samples based on Bray-Curtis (a) and Jensen-Shannon (b) distance methods. Each green (restriction fecal samples) and red (baseline fecal samples) sphere represent the species composition of the entire microbiome present in a given fecal sample.

3.3.4 Microbial community composition of baseline and restriction fecal samples

In order to assess the effect of water restriction on the fecal microbial community composition, we built taxonomic abundance profiles of the 1342 features (ASVs) that remained post data filtration. The relative abundances were summarized at phylum and genus levels.

As depicted in Figure: 3.5(a) and Table: 3.3, out of the 10 most abundant phyla identified from baseline samples, Firmicutes (76.4%), Bacteroidota (13.8%), Euryarchaeota (4.5%), Actinobacteriota (2.4%), Spirochaetota (1.6%), accounted for 97.1% of the total abundance. The rest of the 10 most abundant phyla were Proteobacteria, Verrucomicrobiota, Desulfobacterota, Fibrobacterota, and Cyanobacteria, accounting for 1.19% of the total abundance. For the restriction fecal samples, the list of the 10 most abundant taxa remained the same. However, the relative abundances of some taxa differed significantly.

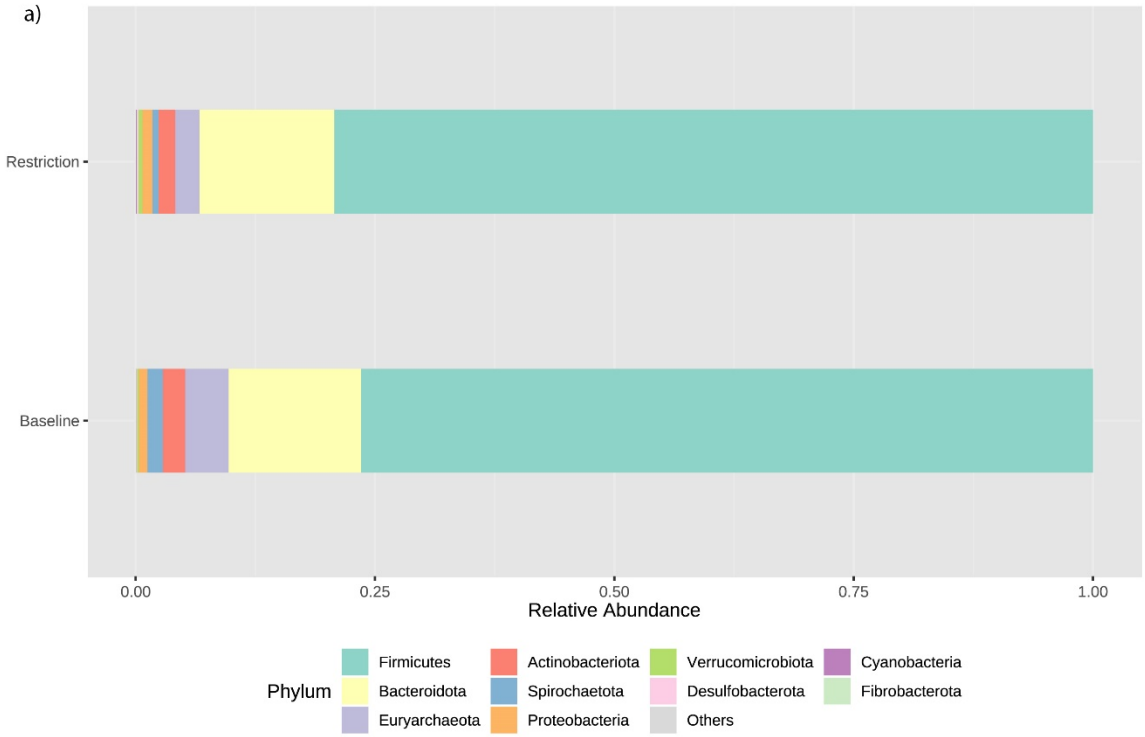
As depicted in Table 3.3, for Firmicutes, Verrucomicrobiota, Desulfobacterota, and Cyanobacteria, we observed a significant increase ($p<0.05$) in their relative abundance in restriction fecal samples compared to baseline. On the other hand, for the phyla Euryarchaeota, Actinobacteriota, and Spirochaetota, we observed significant decrease in their overall relative abundance in restriction fecal samples when compared to baseline. No significant differences ($p<0.05$) were observed in the overall relative abundance of

Bacteroidota, Proteobacteria, and Fibrobacterota between baseline and restriction fecal samples.

We summarized the microbial composition of both baseline and restriction fecal samples at genus level as illustrated in Figure 3.5(b). The 10 most abundant genera for both baseline and restriction fecal samples did not differ in their taxonomy, and it comprised of; *UCG_005, Lachnospiraceae_NK3A20_group, Turicibacter, Clostridium_sensu_stricto_1, Christensenellaceae_R_7_group, Methanobrevibacter, Romboutsia, Rikenellaceae_RC9_gut_group, Ruminococcus*, and *Paeniclostridium*. UCG-005 that belongs to the *Oscillospiraceae* family and Firmicutes had the highest abundance in both baseline (10.0%) and restriction (8.6%) samples.

However, we observed significant variations in the relative abundances of some of the 10 most predominant genera, when an animal wise paired t-test was conducted for each of the top 10 taxa. *Turicibacter, Clostridium_sensu_stricto_1, Christensenellaceae_R_7_group, Romboutsia,* and *Paeniclostridium* had a significant increase ($p<0.05$) in their relative abundances in the restriction samples, compared to baseline. *UCG_005, Methanobrevibacter,* and *Ruminococcus* demonstrated significant decreases ($p<0.05$) in their relative abundances in the restriction samples, compared to baseline. We could not detect any significant differences in the abundances of *Lachnospiraceae_NK3A20_group* and *Rikenellaceae_RC9_gut_group*. 18.3% of the total abundance could not be assigned a taxonomic classification at genus level. The

taxonomic composition of the 20 most abundant genera for both baseline and restriction samples are further elaborated in in Figure 3.6 and Table 3.5.



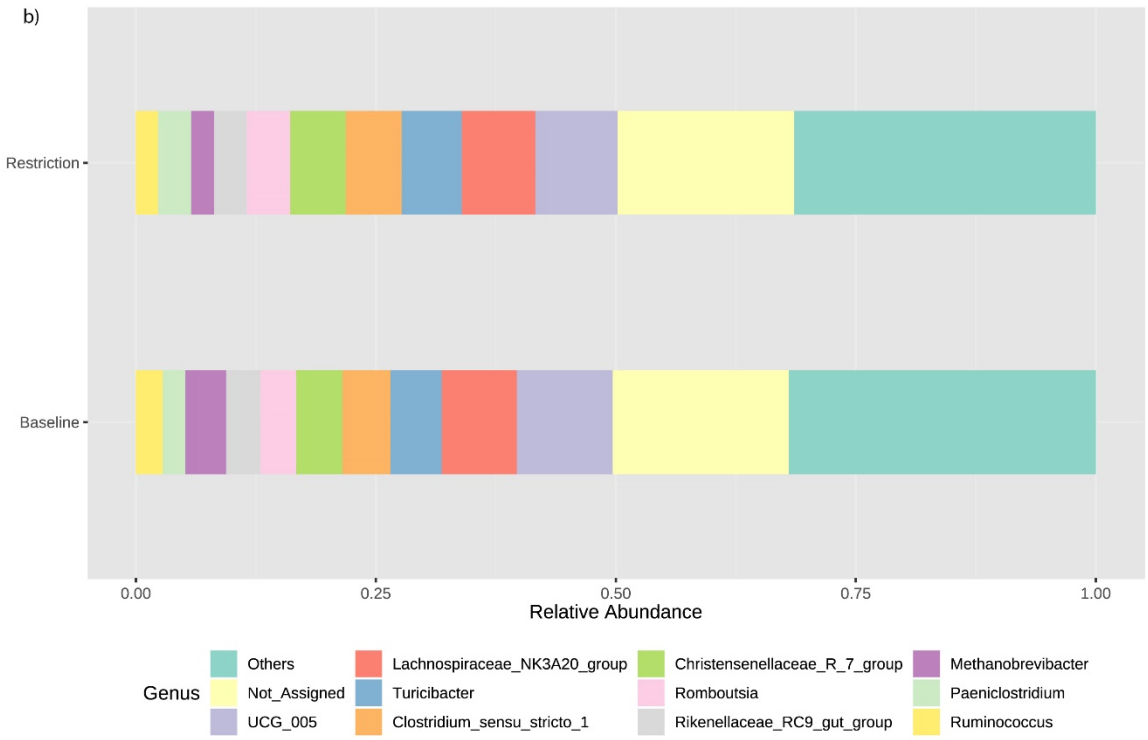


Figure 3.5: Relative abundance bar graphs depicting the 10 most abundant phyla (a) and genera (b) found in baseline and restriction fecal samples.

Phylum	Baseline	Restriction	P-value
Firmicutes	76.44%	79.25%	4.38e-06
Bacteroidota	13.85%	14.09%	8.50e-01
Euryarchaeota	4.53%	2.53%	4.35e-09
Actinobacteriota	2.37%	1.74%	5.77e-06
Spirochaetota	1.57%	0.67%	6.17e-06
Proteobacteria	0.91%	1.04%	1.19e-01
Verrucomicrobiota	0.17%	0.40%	9.58e-04
Desulfobacterota	0.05%	0.11%	1.13e-05
Fibrobacterota	0.04%	0.04%	2.93e-01
Cyanobacteria	0.02%	0.06%	2.74e-07

Table 3.3: Relative abundances of the 10 most abundant phyla in baseline and restriction fecal DNA samples, sorted in descending order of baseline abundances.

Genus	Baseline	Restriction	P-value
UCG_005	10.0%	8.6%	3.79e-05
Lachnospiraceae_NK3A20_group	7.8%	7.7%	0.18
Turicibacter	5.3%	6.3%	3.29e-4
Clostridium_sensu_stricto_1	5.0%	5.8%	4.04e-3
Christensenellaceae_R_7_group	4.8%	5.8%	4.77e-10
Methanobrevibacter	4.3%	2.4%	4.13e-09
Romboutsia	3.7%	4.5%	2.81e-07
Rikenellaceae_RC9_gut_group	3.6%	3.4%	0.23
Ruminococcus	2.8%	2.3%	1.03e-4
Paeniclostridium	2.3%	3.5%	2.76e-14

Table 3. 4: Relative abundances of the 10 most abundant genera in baseline and restriction fecal DNA samples, sorted in descending order of baseline abundances.

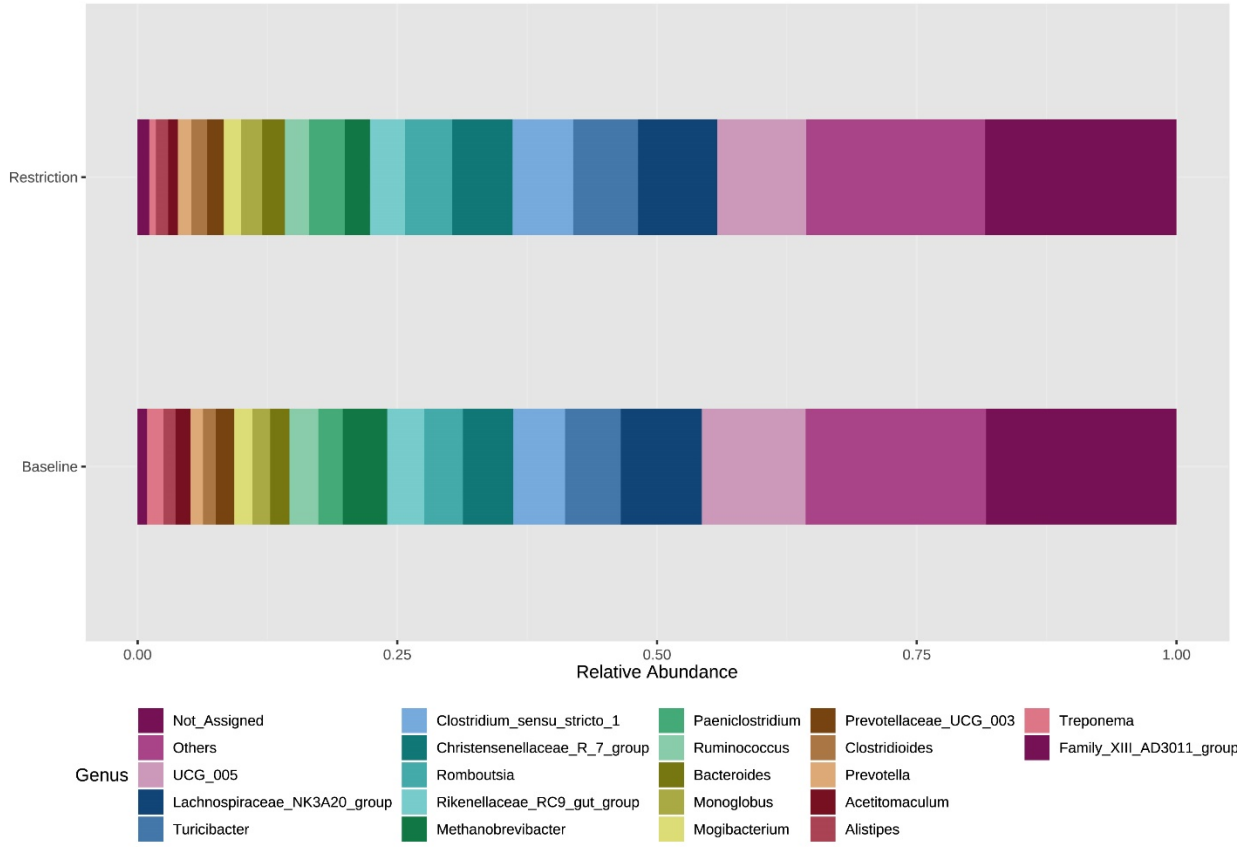


Figure 3.6: The relative abundance bar graphs for restriction and bassline fecal samples at genus level depicting the 20 most abundant genera. The two treatment wise bar graphs are derived by summarizing animal wise microbial abundances.

Genus	Baseline	Restriction
Not_Assigned	18.3%	18.4%
Others	17.4%	17.2%
UCG_005	10.0%	8.6%
Lachnospiraceae_NK3A20_group	7.8%	7.6%
Turicibacter	5.4%	6.2%
Clostridium_sensu_stricto_1	5.0%	5.8%
Christensenellaceae_R_7_group	4.8%	5.8%
Methanobrevibacter	4.3%	2.4%
Romboutsia	3.7%	4.5%
Rikenellaceae_RC9_gut_group	3.6%	3.4%
Ruminococcus	2.8%	2.3%
Paeniclostridium	2.4%	3.4%
Bacteroides	1.9%	2.2%
Mogibacterium	1.8%	1.7%
Prevotellaceae_UCG_003	1.8%	1.6%
Monoglobus	1.7%	2.0%
Treponema	1.6%	0.6%
Acetitomaculum	1.4%	0.9%
Clostridioides	1.3%	1.5%
Prevotella	1.2%	1.3%
Alistipes	1.2%	1.2%
Family_XIII_AD3011_group	0.9%	1.1%

Table 3.5: The relative abundances of the 20 most abundant microbial genera in baseline and restriction fecal samples, sorted in descending of baseline abundances. ~18% of the total sequence abundance remains unclassified at genus level.

3.3.5 The core-microbiome mostly remains stagnant across treatments

To further elucidate the microbial community structure of baseline and restriction fecal samples, we conducted a core microbiome analysis at genus level. The core microbiome was defined as the taxa that are present in at least 70% of the samples considered, at a threshold relative abundance of 0.01% or above.

The core microbiome for baseline fecal samples (Figure 3.7 (a)) consisted of 13 genera, comprising of; *UCG_005, Ruminococcus, Lachnospiraceae_NK3A20_group, Clostridium_sensu_stricto_1, Christensenellaceae_R_7_group, Turicibacter, Rikenellaceae_RC9_gut_group, Mogibacterium, Romboutsia, Methanobrevibacter, Paeniclostridium, Monoglobus, and Bacteroides.*

The core microbiome for fecal samples from the restriction trial (Figure 3.7(b)) included 15 genera comprising of; *UCG_005, Turicibacter, Romboutsia, Lachnospiraceae_NK3A20_group, Clostridium_sensu_stricto_1, Christensenellaceae_R_7_group, Paeniclostridium, Rikenellaceae_RC9_gut_group, Ruminococcus, Mogibacterium, Monoglobus, Clostridioides, Bacteroides, Methanobrevibacter, and Prevotellaceae_UCG_003.*

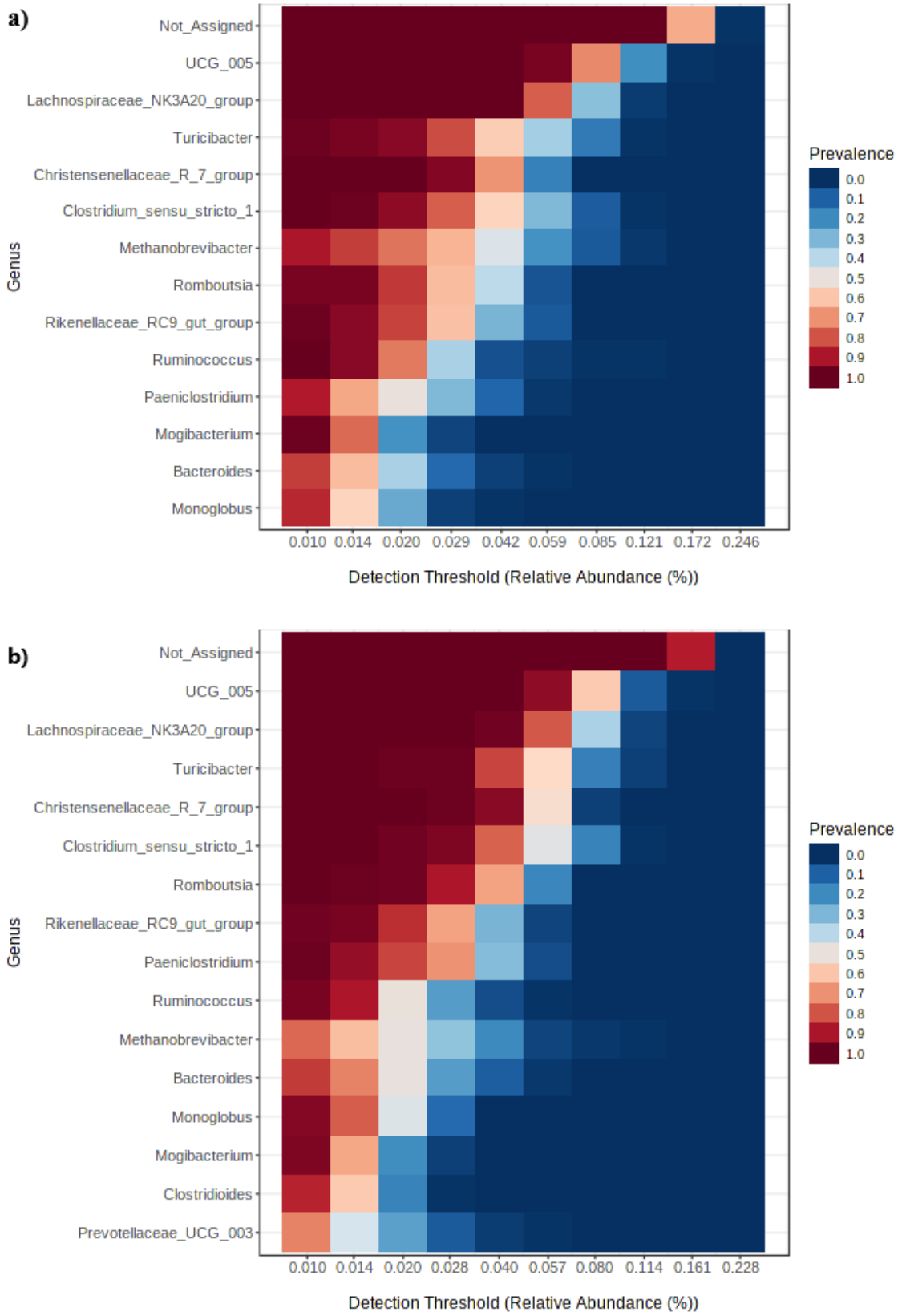

a)
Not_Assigned
UCG_005
Lachnospiraceae_NK3A20_group
Turicibacter
Christensenellaceae_R_7_group
Clostridium_sensu_stricto_1
Methanobrevibacter
Romboutsia
Rikenellaceae_RC9_gut_group
Ruminococcus
Paeniclostridium
Mogibacterium
Bacteroides
Monoglobus
Genus
0.010 0.014 0.020 0.029 0.042 0.059 0.085 0.121 0.172 0.246
Detection Threshold (Relative Abundance (%))
Prevalence
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
b)
Not_Assigned
UCG_005
Lachnospiraceae_NK3A20_group
Turicibacter
Christensenellaceae_R_7_group
Clostridium_sensu_stricto_1
Romboutsia
Rikenellaceae_RC9_gut_group
Paeniclostridium
Ruminococcus
Methanobrevibacter
Bacteroides
Monoglobus
Mogibacterium
Clostridioides
Prevotellaceae_UCG_003
Genus
0.010 0.014 0.020 0.028 0.040 0.057 0.080 0.114 0.161 0.228
Detection Threshold (Relative Abundance (%))
Prevalence
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Figure 3.7: The heat map illustrating the relative abundance and prevalence of the microbial taxa defining the core microbiomes recovered from baseline (a) and restriction (b) fecal samples. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria for microbial taxa to be considered as a part of the baseline or restriction fecal core microbiome.

3.3.6 Biomarker analysis of baseline and restriction fecal microbial populations

In order to identify microbial biomarkers that are enriched in baseline and restriction fecal samples, a LEfSe analysis (Segata et al., 2011) was conducted. We were able to identify 75 differentially abundant taxa between baseline and restriction fecal samples, based on FDR adj. P-value < 0.05 and LDA > 2.0.

As depicted in Figure 3.8, the 15 most influential biomarkers identified by LEfSe, were; *Methanobrevibacter, UCG_005, Paeniclostridium, Christensenellaceae_R_7_group, Treponema, Turicibacter, Clostridium_sensu_stricto_1, Romboutsia, Acetitomaculum, Ruminococcus*, *Cellulosilyticum, UCG_002, Monoglobus, Clostridioides*, and *Coprococcus*.

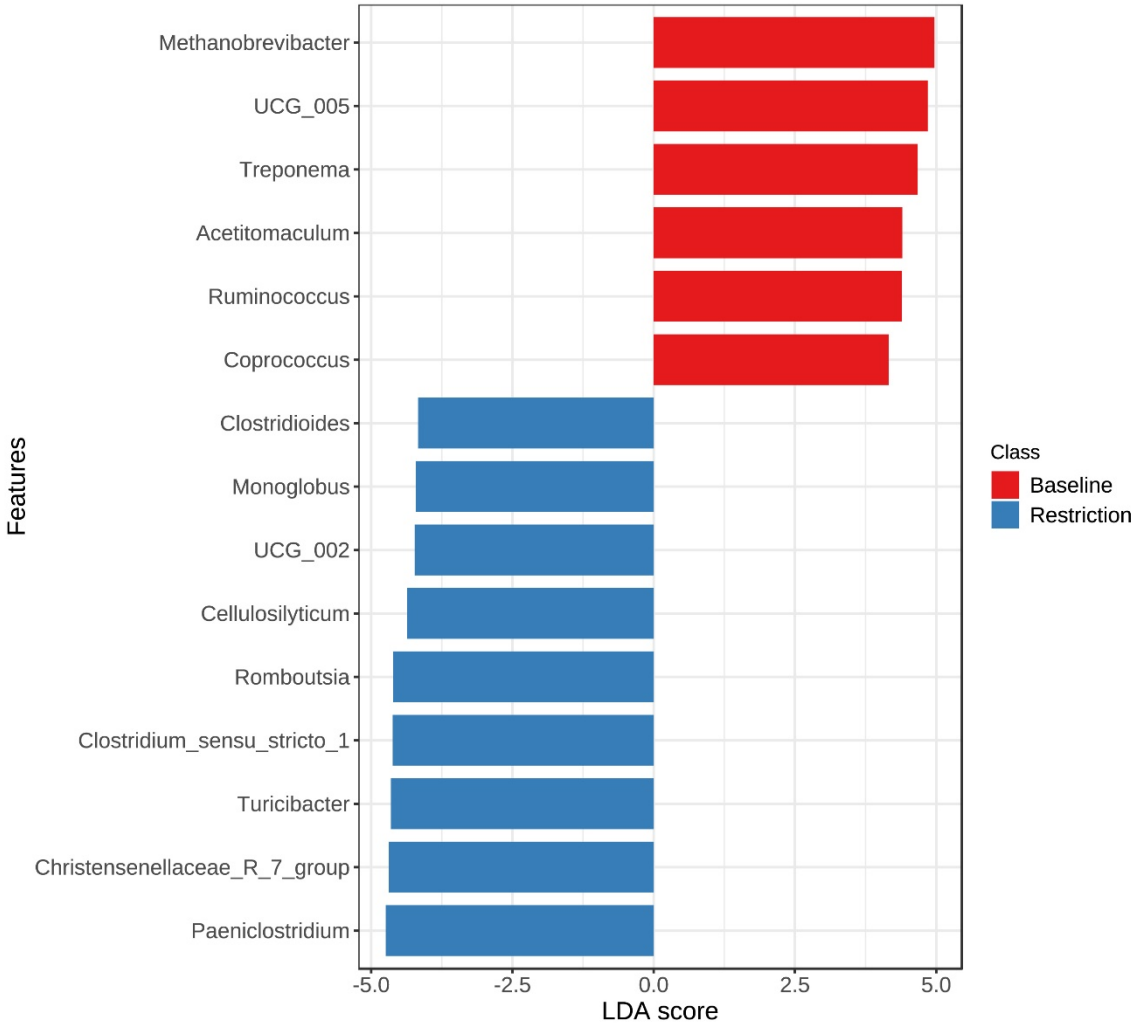


Figure 3.8: The bar plot depicting the 15 most enriched biomarkers ranked in the descending order of their LDA score (magnitude), out of the 75 total biomarkers identified by LEfSe analysis based on FDR adj. P-value < 0.05 and LDA > 2.0. The blue bars indicate the LDA score of biomarkers enriched in restriction fecal samples and the red bars indicate LDA scores of biomarkers enriched in baseline fecal samples. (Segata et al., 2011; Chong et al., 2020).

Random forest (RF) analysis (Breiman, 2001) implemented in MicrobiomeAnalyst (Chong et al., 2020) was used to further elucidate the differentially abundant taxa

between baseline and restriction fecal samples, at genus level. The 15 most significant biomarkers (based on their mean decrease in accuracy) identified by RF analysis (depicted in Figure: 3.9) were; UCG_002, *Cellulosilyticum, Marvinbryantia, UCG_007, Anaerovibrio, Christensenellaceae_R_7_group, Agathobacter, dgA_11_gut_group, Syntrophococcus, Acetitomaculum, Paeniclostridium, Treponema, p_1088_a5_gut_group, Family_XIII_AD3011_group,* and *Methanobrevibacter.*

Genera such as *Marvinbryantia, Agathobacter, Syntrophococcus, Acetitomaculum, Treponema,* and *Methanobrevibacter* were associated with baseline fecal samples, and *UCG_002, Cellulosilyticum, UCG_007, Anaerovibrio, Christensenellaceae_R_7_group, dgA_11_gut_group, Paeniclostridium, p_1088_a5_gut_group,* and *Family_XIII_AD3011_group* were associated with restriction fecal samples.

Only 7 out of the 15 biomarkers identified by the RF analysis overlapped with the 15 most significant biomarkers identified by LEfSe analysis. However, all 15 overlapped with the 75 differentially abundant taxa (LDA>2.0 & FDR-adj. P-value<0.05) identified by LEfSe analysis.

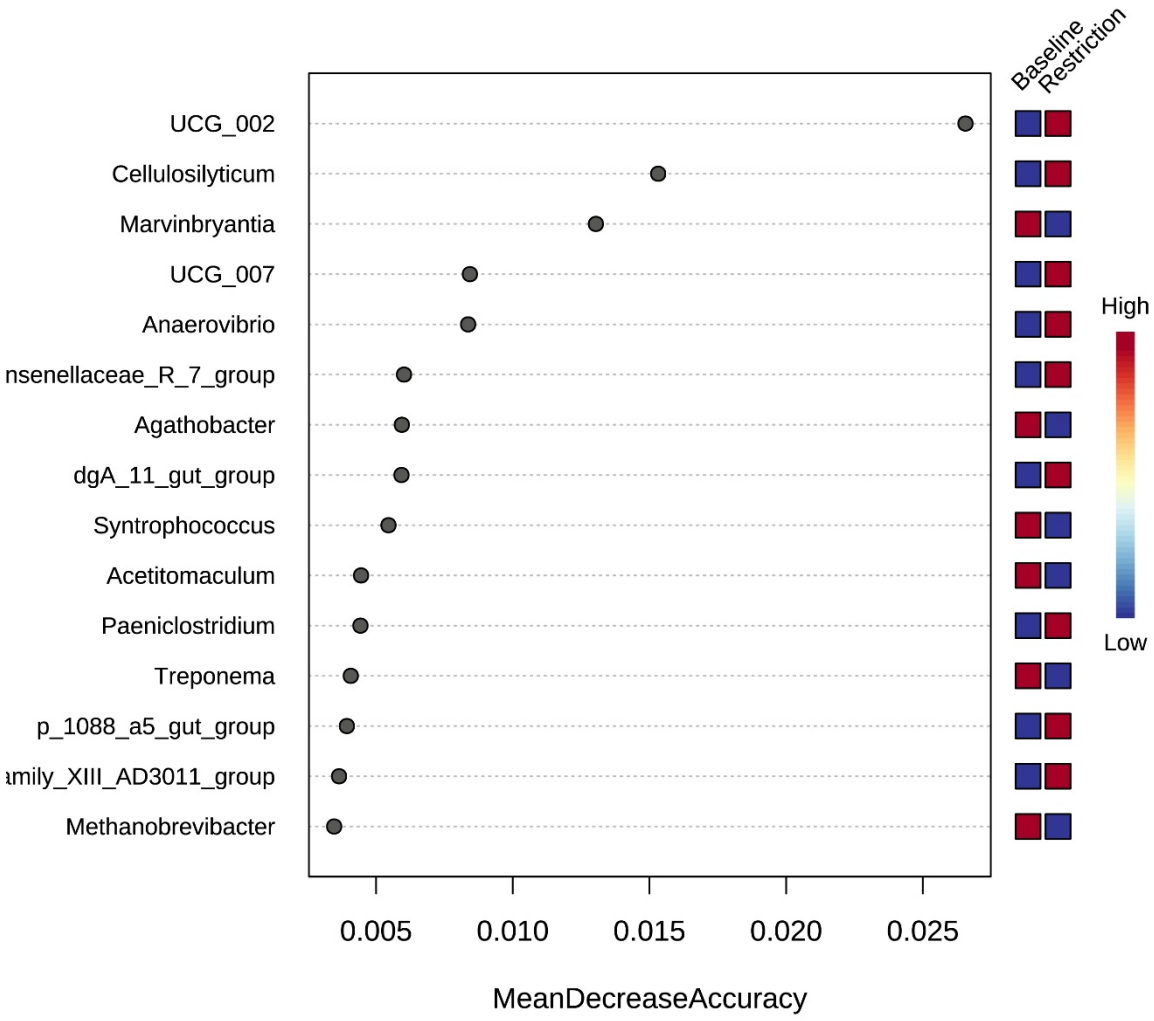


Figure 3.9: Illustration of the differentially abundant features identified by random forest analysis. The features are ranked according to the mean decrease in accuracy the model suffers in assigning samples to the classes, if the particular feature is removed from the analysis. The color codes indicate the association of each feature either with baseline or restriction samples.

Genera	P-values	FDR-adj. P-value	LDAscore
*Methanobrevibacter**	3.72e-8	3.02e-7	4.97
UCG_005	3.86e-6	1.88e-5	4.85
*Paeniclostridium**	8.45e-10	1.03e-8	-4.74
*Christensenellaceae_R_7_group**	4.29e-10	5.69e-9	-4.69
*Treponema**	1.20e-7	7.30e-7	4.67
Turicibacter	8.63e-4	2.248e-3	-4.66
Clostridium_sensu_stricto_1	1.27e-3	3.20e-3	-4.62
Romboutsia	2.19e-5	7.61e-5	-4.61
*Acetitomaculum**	4.36e-11	7.96e-10	4.4
Ruminococcus	1.55e-4	4.61e-4	4.38
*Cellulosilyticum**	3.20e-17	1.56e-15	-4.36
*UCG_002**	7.34e-22	1.07e-19	-4.24
Monoglobus	6.24e-5	2.07e-4	-4.22
Clostridioides	9.28e-6	3.66e-05	-4.17
Coprococcus	5.22e-9	5.44e-08	4.16

(*) – Overlap with top 15 biomarkers from RF analysis

Table 3.6: The top 15 biomarkers identified during LEfSe analysis, and their overlap (indicated using *) with biomarkers identified by RF analysis. The unmarked biomarkers were uniquely identified by the LEfSe analysis.

3.4 Discussion

The GI tract microbiota plays an essential role in the cattle industry due to its influence on animal health, well-being, performance and impact on the environment. Thus, it becomes imperative to have a comprehensive understanding about these complex and dynamic microbial communities. Hence the necessity to generate the knowledge required to better inform livestock producers in decision making with regard to health and well-being of animals, and sustainable growth of their cattle operations (Lourenco et al., 2020).

Even though a multitude of studies based on culture independent methods have focused on elucidating the dynamics of the bovine GI tract microbiome, a vast majority of that have focused on the rumen microbial population. Hence, information about the lower gut microbial communities in cattle is limited (Lourenco et al., 2020). Based on existing literature, diet composition is one of the most influential as well as extensively studied factors that affect the microbial dynamics of the bovine fecal microbiome (Shanks et al., 2011; Cendron, 2020). More specifically the forage to concentrate ratio in the diet provided for feedlot cattle is considered a major aspect of the dietary influence on the fecal microbiome (Kim et al., 2014). However, the effect of factors such water intake, water restriction, water temperature, pH, and dissolved mineral content on the bovine fecal microbiome is yet to be studied extensively.

Therefore, in order to assess the influence of stress created by water restricted conditions on the fecal microbial dynamics, we reconstructed microbial communities in fecal samples obtained from animals maintained at 50% water restriction (restriction fecal samples). These microbial communities derived using 16S rRNA V4 gene amplicon sequencing were then compared against similarly reconstructed microbial communities of fecal samples obtained from the same animals under *ad libitum* water intake.

3.4.1 Alpha diversity measures indicate a significant increase in species evenness in fecal samples obtained during water restriction.

Based on the results of the species richness and evenness indices, it can be concluded that the alpha diversity of the microbial communities recovered from the restriction fecal samples have increased, when compared to baseline samples. However, the indices that take into account the combinatorial effect of species richness and evenness such as Shannon and Simpson's, were the only ones that were able to capture this. Indices that take into account only one aspect of alpha diversity, either species richness (Chao1 and Observed ASV indices) or evenness (Fisher's index) may not have been sensitive enough to capture the true between group variation in alpha diversity between baseline and restriction rumen samples.

3.4.2 Fecal samples obtained during *ad libitum* water intake, and water restriction differ significantly in their species composition.

Moderate clustering with clear overlap was observed between restriction and baseline fecal samples for both Bray-Curtis index and Jensen-Shannon divergence. Furthermore, between group comparisons for restriction and baseline fecal samples, using PERMANOVA revealed high statistical significance (p-value < 0.05). Microbial communities recovered from fecal samples when the animal was under water stress are significantly different in their beta diversity (species composition), when compared against fecal microbial communities recovered at *ad libitum* water intake.

3.4.3 The taxonomy of prominent fecal microbiota remains unchanged during water restriction. Their relative abundances, however, differ significantly.

Next generation sequencing data from microbiome studies are always constrained by the sequencing depth of the NGS platform used. Therefore, unlike a typical ecological study, the abundance of a particular species cannot be considered independent of the abundance of its co-inhabitants. Due to this inherent compositionality of microbiome NGS data, taxonomic abundance is best depicted using relative abundance values (Gloor et al., 2017). Hence, we summarized data at genus and phylum levels to generate relative abundance graphs and tables to study microbial community composition of fecal samples.

We were able to identify Firmicutes, Bacteroidota, Euryarchaeota, Actinobacteriota, Spirochaetota, Proteobacteria, Verrucomicrobiota, Desulfobacterota, Fibrobacterota, and Cyanobacteria, as the top 10 phyla for both baseline and restriction fecal samples. Even though the taxonomy of the 10 most abundant phyla remained the same for both baseline and restriction fecal samples, there were significant differences ($p<0.05$) observed in some of their relative abundances.

Our findings for 10 most prominent phyla fall mostly in line with existing literature. Similar to our findings, Firmicutes and Bacteroidota (Bacteroidetes) have been described the most dominant phyla in the cattle fecal microbiome during large cohort studies (Hagey et al., 2019). Other than that, Euryarchaeota, Actinobacteriota (Actinobacteria), Spirochaetota (Spirochetes), Proteobacteria, Verrucomicrobiota (Verrucomicrobia), and Fibrobacterota (Fibrobacter) have also been found to be common dominant inhabitants of

the cattle fecal microbiome (Shanks et al., 2011; Mote et al., 2019; Lourenco et al., 2020).

Desulfobacterota however is not commonly observed as a dominant phylum in the fecal or rumen microbiome of cattle. The Desulfobacterota identified in our study were largely unclassified at genus level, other than for two genera; *Desulfovibrio* and *Mailhella*. *Desulfovibrio* is a gram-negative sulfate reducing bacteria (SRB) commonly found in the rumen of cattle fed with high concentrate and by-products of the ethanol industry. SRB play a pivotal role in reducing the concentrations of sulfate that can be toxic to the animal, by reducing it to H_2S, which is subsequently eructed out (Richter, 2011).

Methanogens such as *Methanobrevibacter* utilize hydrogen produced during the rumen fermentation process to reduce CO_2 into methane. Similarly, rumen *Desulfovibrio,* use hydrogen to make sulfide from sulfate. The existing literature suggest a positive correlation between *Desulfovibrio* and high methane emissions in cattle, since high emissions may imply high H_2 concentration inside the rumen thus facilitating the growth of *Desulfovibrio* (Wallace et al., 2015). However, in the current study we observed a decrease in the abundance of *Methanobrevibacter* while an increase in abundance of *Desulfovibrio* in restriction samples. This discrepancy may still be explained by the fact that in a situation where two H_2 utilizers compete for the available hydrogen in the rumen (Shah et al., 2020), one H_2 utilizer (*Desulfovibrio)* may thrive in the absence of another (*Methanobrevibacter).*

Mailhella on the other hand, has been observed in the fecal microbiome of heifers, although information on its potential role in the bovine gut microbiome is still not available (Cendron, 2020). The relative abundance of both *Desulfovibrio* and *Mailhella* increased in restriction fecal samples (*Desulfovibrio* - 0.03%, *Mailhella* – 0.09%), compared to baseline samples (*Desulfovibrio* - 0.02%, *Mailhella* - 0.03%), even though their overall abundance was very low.

When the taxonomic abundance data were summarized at genus level, again we observed that the ten most dominant genera remained the same for both baseline and restriction fecal samples. However, the relative abundance for many of them differed significantly. The predominant genera identified by us have been previously described in the studies concerning the fecal microbiota of cattle (Cendron, 2020; Huang et al., 2021). Hence, our findings are in agreement with existing knowledge on the composition of fecal microbiome of cattle.

In comparison to baseline samples, the restriction fecal samples had significant increases in the relative abundance of *Turicibacter*, *Clostridium_sensu_stricto_1*, *Christensenellaceae_R_7_group*, *Romboutsia*, and *Paeniclostridium*. *Clostridium_sensu_stricto_1* is a genus commonly associated with pathogenic bacteria (Huang et al., 2021). *Christensenella minuta* is the species most closely related to the *Christensenellaceae_R_7_group* and is known to be associated with pH reduction in the rumen. *C. minuta* is also known to produce acetate and butyrate thorough glucose metabolism (Holman and Gzyl, 2019). The genus *Turicibacter* is known to be associated

with feed intake and has been detected in human gut microbiome as well. However, still very little information is available for this genus, hence further studies are much warranted especially due to animal gut health implications since some species in this genera are considered pathogenic (P R Myer et al., 2015). Both *Romboutsia* and *Paeniclostridium* are known to be associated with diet composition (Cendron, 2020). *Romboutsia* has also found to be the among the most abundant microbiota in the small intestine of Nelore steers (Lopes et al., 2019).

The increase of abundance in phylum Verrucomicrobiota and subsequent increase in the abundance of the genus *Akkermansia* in restriction fecal samples is also noteworthy. *Akkermansia* are known to be mucine-degrading bacteria and have been observed to be associated with Angus cattle. Mucine is an important component in the mucosal barrier of the gut, and mucine degradation can result in increased susceptibility of the host to gut pathogens (Fan et al., 2020). Hence, the increase is relative abundance of *Akkemansia* we observed in restriction fecal samples can be considered as further evidence of adverse gut health implications due to water restricted conditions.

3.4.4 The fecal core-microbiome remains essentially unchanged during water restriction, with exception of an addition of pathogenic genus *Clostridioides*

The core microbiome is a contingent of microbial taxa that usually plays an essential role in host-microbiome interactions, hence found to be conserved across time, breed, treatment and sppatial distribution. However, due to lack of consensus, a clear definition on parameters and their threshold values to define a core-microbiome for a given

environment is yet to materialize (Risely, 2020). The prevalence of an organism across hosts or samples, and its relative abundance in the microbial communities recovered, are commonly used as the parameters to define a core microbiome. Their thresholds however, may defer from study to study (Wallace et al., 2019).

A prevalence threshold of 70% and relative abundance cut-off value of 0.01% was used to define the core microbiome at genus level for both baseline and restriction core microbiomes for the fecal samples of this study. We were able to identify a core microbiome comprising of 13 genera for baseline fecal samples, and a core microbiome of 15 genera for restriction fecal samples. The core microbiome for restriction fecal samples essentially contained the 13 genera found in baseline fecal core microbiome, with the addition of *Clostridioides* and *Prevotellaceae_UCG_003*.

Clostridioides is a reclassified genus containing taxa that previously belonged to *Clostridium,* and contains only two child taxa, namely; *Clostridioides difficile* and *Clostridioides mangenotii* (Lawson et al., 2016).

Clostridioides difficile is described as a Gram-positive, anaerobic bacterium that can cause severe colitis or *C. difficile* infection (CDI) in humans. It has also been observed in the gut and feces of animals such as swine, horses and cattle. The zoonotic transmission of CDI happens through the fecal-oral route. *C. difficile* has been observed in both adult dairy cows and beef calves (Redding et al., 2021). However, the exact health implications of *C. difficile* in cattle is still unclear and is believed to be minimal. There are reports of

an association between *C. difficile* in feces and diarrhea in calves, however, the results could not be experimentally reproduced (Weese, 2020). We also observed an increase in relative abundance of *Clostridioides* in restriction fecal samples (1.54%), compared to the baseline (1.25%).

3.4.5 LEfSe analysis and Random Forest classification reveal differentially abundant taxa (biomarkers) associated with baseline and restriction fecal microbiomes.

To identify biomarkers for baseline and restriction fecal samples, we used an FDR adj. p-value <0.05 and LDA>2.0 with all LEfSe analyses done at genus level (Chong et al., 2020).

In the current study, six out of the 15 differentially abundant taxa identified by LEfSe analysis, *Methanobrevibacter, UCG_005, Treponema, Acetitomaculum*, and *Coprococcus* were highly associated with baseline samples. The other 9 taxa, *Clostridoides, Monoglobus, UCG_002, Cellulosilyticum, Clostridum_sensu_stricto_1, Turicibacter, Christensenellaceae_R_7_group,* and *Pauniclostridum* were highly associated with restriction fecal samples.

UCG-002 that belongs to the *Oscillospiraceae* family, *Monoglobus*, and *Cellulosilyticum* were the genera that did not appear in the list of predominant genera (with increased relative abundances for restriction fecal samples) and the core microbiome for restriction fecal samples. When relative abundance data were mined for these three taxa, it was

observed that their relative abundance increased during water restriction, when compared to baseline. *Cellulosilyticum* – from 0.45% to 0.90%; *Monoglobus* – from 1.69% to 2.02%; *UCG_002* – from 0.23% to 0.58%. Thus making it clear that results generated from different perspectives are in good agreement.

Monoglobus pectinilyticus is the only child taxon described under the *Monoglobus* genus and has been studied as a model organism for pectin degradation in the human gut. OTUs with 95% similarity to *Monoglobus* have been described in cattle fed with Tall fescue grass (Kim et al., 2017; Kim et al., 2019; Koeste et al., 2020). Genus *Cellulosilyticum* comprises of cellulolytic bacteria associated with the cecum of cattle. Increased abundance of *Cellulosilyticum* has been associated with inflammatory responses in aging cows, and with feedlot cattle fed with antibiotics such as monensin and tylosin as feed additives in finishing diets (Thomas et al., 2017; Zhang et al., 2019)

Out of the 15 most influential biomarkers identified by the RF analysis, only 7 overlapped with the 15 most significant biomarkers identified by LEfSe analysis. Even though, all 15 of them overlapped with the 75 differentially abundant taxa identified by LEfSe analysis. *Marvinbryantia, UCG_007,* and *Anaerovibrio* should be noted as biomarkers (with relatively high mean decrease in accuracy) identified by RF analysis, but failed to appear in the list of predominant differentially abundant taxa derived using LEfSe analysis. To measure the magnitude of the effect exerted by a particular biomarker in classifying the samples, RF and LEfSe takes two different approaches. RF uses ‘mean decrease in accuracy’ which is a direct implication of a supervised non-parametric

learning algorithm (Breiman, 2001; Chong et al., 2020). For the same purpose, LEfSe uses LDA scores that are linear and parametric (Segata et al., 2011). Hence the two different approaches taken by the two algorithms in determining the effect of a 'biomarker' in sample classification may well be the reason behind this discrepancy in identifying top biomarkers, or differentially abundant features. However, looking at data from two different perspectives paves the way for robust conclusions. The fact that we ultimately observed complete overlap of biomarkers identified by both LEfSe and RF analysis, increases the confidence in our data, and future research based on them.

3.5 References

Ahlberg, C. M. 2018. Characterization of Water Intake in Beef Cattle: Test Length Guidelines, Water Intake Prediction, and Genetic Parameters. PhD Diss. Kansas State University, Manhattan.

Breiman, L. 2001. Random Forests. Mach. Learn. 45:5–32. doi:10.1023/A:1010933404324.

Bruno, K. A. 2019. Effects of Extended Water Restriction On Performance, Behavior, Health, And Welfare of Feedlot Steers. PhD Diss. Oklahoma State University, Stillwater.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583. doi:10.1038/nmeth.3869.

Castro, J. J., A. Gomez, B. White, J. R. Loften, and J. K. Drackley. 2016. Changes in the intestinal bacterial community , short-chain fatty acid profile , and intestinal development of preweaned Holstein calves . 2 . Effects of gastrointestinal site and age. J. Dairy Sci. 99:9703–9715. doi:10.3168/jds.2016-11007.

Cendron, F., G. Niero, G. Carlino, M. Penasa, and M. Cassandro. 2020. Characterizing the fecal bacteria and archaea community of heifers and lactating cows through 16S rRNA next-generation sequencing. J. Appl. Genet. 61:593–605. doi:10.1007/s13353-020-00575-3.

Chao, A., and L. Jost. 2012. Coverage-based rarefaction and extrapolation: standardizing samples by completeness rather than size. Ecology. 93:2533–2547. doi:10.1890/11-1952.1.

Chong, J., P. Liu, G. Zhou, and J. Xia. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat. Protoc. 15:799–821. doi:10.1038/s41596-019-0264-1.

Clemmons, B. A., S. T. Reese, F. G. Dantas, G. A. Franco, T. P. L. Smith, O. I. Adeyosoye, K. G. Pohler, and P. R. Myer. 2017. Vaginal and uterine bacterial communities in postpartum lactating cows. Front. Microbiol. 8:1–10. doi:10.3389/fmicb.2017.01047.

Cremer, J., M. Arnoldini, and T. Hwa. 2017. Effect of water flow and chemical environment on microbiota growth and composition in the human colon. Proc. Natl. Acad. Sci. U. S. A. 114:6438–6443. doi:10.1073/pnas.1619598114.

Dethlefsen, L., S. Huse, M. L. Sogin, and D. A. Relman. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16s rRNA sequencing.

PLoS Biol. 6:2383–2400. doi:10.1371/journal.pbio.0060280.

Deusch, S., A. Camarinha-Silva, J. Conrad, U. Beifuss, M. Rodehutscord, and J. Seifert. 2017. A structural and functional elucidation of the rumen microbiome influenced by various diets and microenvironments. Front. Microbiol. 8:1605. doi:10.3389/fmicb.2017.01605.

Dhariwal, A., J. Chong, S. Habib, I. L. King, L. B. Agellon, and J. Xia. 2017. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 45:W180–W188. doi:10.1093/nar/gkx295.

Fan, P., B. Bian, L. Teng, C. D. Nelson, J. Driver, M. A. Elzo, and K. C. Jeong. 2020. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 14:302–317. doi:10.1038/s41396-019-0529-2.

Gloor, G. B., J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. 2017. Microbiome datasets are compositional: And this is not optional. Front. Microbiol. 8:1–6. doi:10.3389/fmicb.2017.02224.

Good, I. J., and G. H. Toulmin. 1956. The Number Of New Species, And The Increase In Population Coverage, When A Sample Is Increased. Biometrika. 43:45–63. doi:10.1093/biomet/43.1-2.45.

Gressley, T. F., M. B. Hall, and L. E. Armentano. 2011. Ruminant nutrition symposium: Productivity, digestion, and health responses to hindgut acidosis in ruminants. J. Anim. Sci. 89:1120–1130. doi:10.2527/jas.2010-3460.

Sofi, M. H., R. Gudi, S. Karumuthil-Melethil, N. Perez, B. M. Johnson, and C. Vasu. 2014. pH of Drinking Water Influences the Composition of Gut Microbiome and Type 1 Diabetes Incidence. Diabetes. 63:632–644. doi:10.2337/db13-0981.

Hagey, J. V, S. Bhatnagar, J. M. Heguy, B. M. Karle, P. L. Price, D. Meyer, and E. A. Maga. 2019. Fecal microbial communities in a large representative cohort of California dairy cows. Front. Microbiol. 10:1–14. doi:10.3389/fmicb.2019.01093.

Hara, E. O., A. Kelly, M. S. Mccabe, D. A. Kenny, L. L. Guan, and S. M. Waters. 2018. Effect of a butyrate-fortified milk replacer on gastrointestinal microbiota and products of fermentation in artificially reared dairy calves at weaning. 1–11. doi:10.1038/s41598-018-33122-6.

Henderson, G., F. Cox, S. Ganesh, A. Jonker, W. Young, P. H. Janssen, L. Abecia, E. Angarita, P. Aravena, G. N. Arenas, C. Ariza, G. T. Attwood, J. M. Avila, J. Avila-Stagno, A. Bannink, R. Barahona, M. Batistotti, M. F. Bertelsen, A. Brown-Kav, A. M. Carvajal, L. Cersosimo, A. V. Chaves, J. Church, N. Clipson, M. A. Cobos-Peralta, A. L. Cookson, S. Cravero, O. C. Carballo, K. Crosley, G. Cruz, M. C. Cucchi, R. De La Barra, A. B. De Menezes, E. Detmann, K. Dieho, J. Dijkstra, W. L. S. Dos Reis, M. E. R. Dugan, S. H. Ebrahimi, E. Eythórsdóttir, F. N. Fon, M. Fraga, F. Franco, C. Friedeman, N. Fukuma, D. Gagić, I. Gangnat, D. J. Grilli, L. L. Guan, V. H. Miri, E. Hernandez-Sanabria, A. X. I. Gomez, O. A. Isah, S. Ishaq, E. Jami, J. Jelincic, J. Kantanen, W. J. Kelly, S. H. Kim, A. Klieve, Y. Kobayashi, S. Koike, J. Kopecny, T. N. Kristensen, S. J.

Krizsan, H. LaChance, M. Lachman, W. R. Lamberson, S. Lambie, J. Lassen, S. C. Leahy, S. S. Lee, F. Leiber, E. Lewis, B. Lin, R. Lira, P. Lund, E. Macipe, L. L. Mamuad, H. C. Mantovani, G. A. Marcoppido, C. Márquez, C. Martin, G. Martinez, M. E. Martinez, O. L. Mayorga, T. A. McAllister, C. McSweeney, L. Mestre, E. Minnee, M. Mitsumori, I. Mizrahi, I. Molina, A. Muenger, C. Munoz, B. Murovec, J. Newbold, V. Nsereko, et al. 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5. doi:10.1038/srep14567.

Holman, D. B., and K. E. Gzyl. 2019. A meta-analysis of the bovine gastrointestinal tract microbiota. FEMS Microbiol. Ecol. 95:1–9. doi:10.1093/femsec/fiz072.

Hooper, L. V, D. R. Littman, A. J. Macpherson, and M. P. Program. 2015. Interactions between the microbiota and the immune system. 336:1268–1273. doi:10.1126/science.1223490.Interactions.

Hoover, W. H. 1978. Digestion and Absorption in the Hindgut of Ruminants. J. Anim. Sci. 46:1789–1799. doi:10.2527/jas1978.4661789x.

Huang, S., S. Ji, F. Wang, J. Huang, G. M. Alugongo, and S. Li. 2021. Dynamic changes of the fecal bacterial community in dairy cows during early lactation. AMB Express. doi:10.1186/s13568-020-01106-3.

Kandlikar, G. S., Z. J. Gold, M. C. Cowen, R. S. Meyer, A. C. Freise, N. J. B. Kraft, J. Moberg-Parker, J. Sprague, D. J. Kushner, and E. E. Curd. 2018. Ranacapa: An R package and shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations [version 1; referees: 1 approved, 2 approved with reservations]. F1000Research. 7:1–18. doi:10.12688/f1000research.16680.1.

Kiernan, D. 2020. Introduction, Simpson's Index and Shannon-Weiner Index.

Kim, C. C., G. R. Healey, W. J. Kelly, M. L. Patchett, Z. Jordens, G. W. Tannock, I. M. Sims, T. J. Bell, D. Hedderley, B. Henrissat, and D. I. Rosendale. 2019. Genomic insights from Monoglobus pectinilyticus: a pectin-degrading specialist bacterium in the human colon. ISME J. 13:1437–1456. doi:10.1038/s41396-019-0363-6.

Kim, C. C., W. J. Kelly, M. L. Patchett, G. W. Tannock, Z. Jordens, H. M. Stoklosinski, J. W. Taylor, I. M. Sims, T. J. Bell, and D. I. Rosendale. 2017. Monoglobus pectinilyticus gen. Nov., sp. nov., a pectinolytic bacterium isolated from human faeces. Int. J. Syst. Evol. Microbiol. 67:4992–4998. doi:10.1099/ijsem.0.002395.

Kim, M., J. Kim, L. A. Kuehn, J. L. Bono, E. D. Berry, N. Kalchayanand, H. C. Freetly, A. K. Benson, and J. E. Wells. 2014. Investigation of bacterial diversity in the feces of cattle fed different diets. J. Anim. Sci. 92:683–694. doi:10.2527/jas.2013-6841.

Koeste, L. R., D. H. Pool, N. V. L. Serão, and S. Schmitz-Esser. 2020. Beef cattle that respond differently to fescue toxicosis have distinct gastrointestinal tract microbiota. PLoS One. 15:1–20. doi:10.1371/journal.pone.0229192.

Lawson, P. A., D. M. Citron, K. L. Tyrrell, and S. M. Finegold. 2016. Reclassification of Clostridium difficile as Clostridioides difficile (Hall and O'Toole 1935) Prévot 1938.

Anaerobe. 40:95–99. doi:10.1016/j.anaerobe.2016.06.008.

Lim, M. Y., E. J. Song, S. H. Kim, J. Lee, and Y. Do Nam. 2018. Comparison of DNA extraction methods for human gut microbial community profiling. Syst. Appl. Microbiol. 41:151–157. doi:10.1016/j.syapm.2017.11.008.

Lopes, D. R. G., A. J. La Reau, M. De Souza Duarte, E. Detmann, C. B. P. Bento, M. E. Z. Mercadante, S. F. M. Bonilha, G. Suen, and H. C. Mantovani. 2019. The bacterial and fungal microbiota of nelore steers is dynamic across the gastrointestinal tract and its fecal-associated microbiota is correlated to feed efficiency. Front. Microbiol. 10:1–15. doi:10.3389/fmicb.2019.01263.

Lourenco, J. M., T. J. Kieran, D. S. Seidel, T. C. Glenn, M. F. Da Silveira, T. R. Callaway, and R. L. Stewart. 2020. Comparison of the ruminal and fecal microbiotas in beef calves supplemented or not with concentrate. PLoS One. 15:1–17. doi:10.1371/journal.pone.0231533.

Luz Calle, M. 2019. Statistical analysis of metagenomics data. Genomics and Informatics. 17. doi:10.5808/GI.2019.17.1.e6.

Malmuthuge, N., and L. L. Guan. 2017. Understanding host-microbial interactions in rumen : searching the best opportunity for microbiota manipulation. J. Anim. Sci. Biotechnol. 1–7. doi:10.1186/s40104-016-0135-3.

McMurdie, P. J., and S. Holmes. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One. 8. doi:10.1371/journal.pone.0061217.

McMurdie, P. J., and S. Holmes. 2014. Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. PLoS Comput. Biol. 10. doi:10.1371/journal.pcbi.1003531.

Moore, S. G., A. C. Ericsson, S. E. Poock, P. Melendez, and M. C. Lucy. 2017. Hot topic: 16S rRNA gene sequencing reveals the microbiome of the virgin and pregnant bovine uterus. J. Dairy Sci. 100:4953–4960. doi:10.3168/jds.2017-12592.

Mote, R. S., N. S. Hill, J. H. Skarlupka, Z. B. Turner, Z. P. Sanders, D. P. Jones, G. Suen, and N. M. Filipov. 2019. Response of beef cattle fecal microbiota to grazing on toxic tall fescue. Appl. Environ. Microbiol. 85:1–17. doi:10.1128/AEM.00032-19.

Mulder, I. E., B. Schmidt, M. Lewis, M. Delday, C. R. Stokes, M. Bailey, R. I. Aminov, B. P. Gill, J. R. Pluske, C. D. Mayer, and D. Kelly. 2011. Restricting microbial exposure in early life negates the immune benefits associated with gut colonization in environments of high microbial diversity. PLoS One. 6. doi:10.1371/journal.pone.0028279.

Myer, Phillip R, J. E. Wells, T. P. L. Smith, L. A. Kuehn, and H. C. Freetly. 2015. Microbial community profiles of the colon from steers differing in feed efficiency. Springerplus. doi:10.1186/s40064-015-1201-6.

Myer, P R, J. E. Wells, T. P. L. Smith, L. A. Kuehn, and H. C. Freetly. 2015. Cecum microbial communities from steers differing in feed efficiency. J. Anim. Sci. 93:5327–

5340. doi:10.2527/jas.2015-9415.

Noel, S. J., D. W. Olijhoek, F. Mclean, P. Lovendahl, P. Lund, and O. Hojberg. 2019. Rumen and Fecal Microbial Community Structure of Holstein and Jersey Dairy Cows as Affected by Breed, Diet, and Residual Feed Intake. Animals. 9.

O'Hara, E., A. L. A. Neves, Y. Song, and L. L. Guan. 2020. The Role of the Gut Microbiome in Cattle Production and Health: Driver or Passenger? Annu. Rev. Anim. Biosci. 8:199–220. doi:10.1146/annurev-animal-021419-083952.

Ocejo, M., B. Oporto, and A. Hurtado. 2019. 16S rRNA amplicon sequencing characterization of caecal microbiome composition of broilers and free-range slow-growing chickens throughout their productive lifespan. Sci. Rep. 9:2506. doi:10.1038/s41598-019-39323-x.

Parsons, B. N., U. Z. Ijaz, R. D'Amore, M. D. Burkitt, R. Eccles, L. Lenzi, C. A. Duckworth, A. R. Moore, L. Tiszlavicz, A. Varro, N. Hall, and D. M. Pritchard. 2017. Comparison of the human gastric microbiota in hypochlorhydric states arising as a result of Helicobacter pylori-induced atrophic gastritis, autoimmune atrophic gastritis and proton pump inhibitor use. bioRxiv. 1–18. doi:10.1101/144907.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41:590–596. doi:10.1093/nar/gks1219.

Ramette, A. 2007. Multivariate analyses in microbial ecology. FEMS Microbiol. Ecol. 62:142–160. doi:10.1111/j.1574-6941.2007.00375.x.

Redding, L., E. Huang, J. Ryave, T. Webb, D. Barnhart, L. Baker, J. Bender, M. Kristula, and D. Kelly. 2021. Clostridioides difficile on dairy farms and potential risk to dairy farm workers. Anaerobe. 69:102353. doi:10.1016/j.anaerobe.2021.102353.

Richter, E. L. 2011. The effect of dietary sulfur on performance, mineral status, rumen hydrogen sulfide, and rumen microbial populations in yearling beef steers. MSc Diss. Iowa State University. Ames, Iowa.

Risely, A. 2020. Applying the core microbiome to understand host–microbe systems. J. Anim. Ecol. 89:1549–1558. doi:10.1111/1365-2656.13229.

Roguet, A., A. M. Eren, R. J. Newton, and S. L. McLellan. 2018. Fecal source identification using random forest. Microbiome. 6:1–15. doi:10.1186/s40168-018-0568-3.

Sasada, T., T. Hinoi, Y. Saito, T. Adachi, and Y. Takakura. 2015. Chlorinated Water Modulates the Development of Colorectal Tumors with Chromosomal Instability and Gut Microbiota in Apc -Deficient Mice. 1–15. doi:10.1371/journal.pone.0132435.

Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, and C. Huttenhower. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60. doi:10.1186/gb-2011-12-6-r60.

Shah, A. M., J. Ma, Z. Wang, R. Hu, X. Wang, Q. Peng, F. K. Amevor, and N. Goswami. 2020. Production of Hydrogen Sulfide by Fermentation in Rumen and Its Impact on Health and Production of Animals. Processes. 8:1169. doi:10.3390/pr8091169.

Shanks, O. C., C. A. Kelty, S. Archibeque, M. Jenkins, R. J. Newton, S. L. McLellan, S. M. Huse, and M. L. Sogin. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. Appl. Environ. Microbiol. 77:2992–3001. doi:10.1128/AEM.02988-10.

Smith, B. 2011. Optimising Bacterial DNA Extraction from Faecal Samples: Comparison of Three Methods. Open Microbiol. J. 5:14–17. doi:10.2174/1874285801105010014.

Song, Y., N. Malmuthuge, M. A. Steele, and L. L. Guan. 2018. Shift of hindgut microbiota and microbial short chain pre-weaning. FEMS Microbiol. Ecol. 94:1–15. doi:10.1093/femsec/fix179.

Thomas, M., M. Webb, S. Ghimire, A. Blair, K. Olson, G. J. Fenske, A. T. Fonder, J. Christopher-Hennings, D. Brake, and J. Scaria. 2017. Metagenomic characterization of the effect of feed additives on the gut microbiome and antibiotic resistome of feedlot cattle. Sci. Rep. 7:1–13. doi:10.1038/s41598-017-12481-6.

Wagner, B. D., G. K. Grunwald, G. O. Zerbe, S. K. Mikulich-Gilbertson, C. E. Robertson, E. T. Zemanick, and J. K. Harris. 2018. On the use of diversity measures in longitudinal sequencing studies of microbial communities. Front. Microbiol. 9. doi:10.3389/fmicb.2018.01037.

Wallace, R. J., J. A. Rooke, N. McKain, C. A. Duthie, J. J. Hyslop, D. W. Ross, A. Waterhouse, M. Watson, and R. Roehe. 2015. The rumen microbial metagenome associated with high methane production in cattle. BMC Genomics. 16:839. doi:10.1186/s12864-015-2032-0.

Wallace, R. J., G. Sasson, P. C. Garnsworthy, I. Tapio, E. Gregson, P. Bani, P. Huhtanen, A. R. Bayat, F. Strozzi, F. Biscarini, T. J. Snelling, N. Saunders, S. L. Potterton, J. Craigon, A. Minuti, E. Trevisi, M. L. Callegari, F. P. Cappelli, E. H. Cabezas-Garcia, J. Vilkki, C. Pinares-Patino, K. O. Fliegerová, J. Mrázek, H. Sechovcová, J. Kopečný, A. Bonin, F. Boyer, P. Taberlet, F. Kokou, E. Halperin, J. L. Williams, K. J. Shingfield, and I. Mizrahi. 2019. A heritable subset of the core rumen microbiome dictates dairy cow productivity and emissions. Sci. Adv. 5:eaav8391. doi:10.1126/sciadv.aav8391.

Weese, J. S. 2020. Clostridium (Clostridioides) difficile in animals. J. Vet. Diagnostic Investig. 32:213–221. doi:10.1177/1040638719899081.

Zhang, G., Yachun Wang, H. Luo, W. Qiu, H. Zhang, L. Hu, Yajing Wang, G. Dong, and G. Guo. 2019. The association between inflammaging and age-related changes in the ruminal and fecal microbiota among lactating holstein cows. Front. Microbiol. 10:1–17. doi:10.3389/fmicb.2019.01803.

Zhou, M., Y. J. Peng, Y. Chen, C. M. Klinger, M. Oba, J. X. Liu, and L. L. Guan. 2018. Assessment of microbiome changes after rumen transfaunation: implications on improving feed efficiency in beef cattle. Microbiome. 6:62. doi:10.1186/s40168-018-0447-y.

CHAPTER IV

ASSOCIATION BETWEEN ANIMAL PERFORMANCE AND THE POPULATION DYNAMICS OF THE GUT MICROBIOME OF FEEDLOT CATTLE UNDER WATER RESTRICTED CONDITIONS

4.1 Introduction

The gut microbiome of cattle, commonly represented by its rumen and fecal microbiota, is known to be largely affected by diet. Factors such as host genetics, breed, age, and the external environment are also known to influence its composition (Holman and Gzyl, 2019). Water is considered a key component of an animal's diet that contributes directly or indirectly to almost all of its essential physiological functions connected to digestion, growth, maintenance, fattening, lactation and pregnancy (Lardner et al., 2005; Williams et al., 2017).

The main source of water for most cattle is drinking water. Factors such as impurities in drinking water, water temperature and restricted access to drinking water can greatly decrease water intake, thus decreasing feed intake and ultimately having a profound impact on animal performance (Lofgreen et al., 1975; Wright, 2007). Furthermore, water treatment, coagulation, aeration and overall increase in palatability of drinking water has been linked to a ~10% weight gain in grazing cattle (Lardner et al., 2005), thus providing further evidence for the association between water intake/quality and performance in cattle.

The microbial community residing in both the foregut (rumen) and the lower gastrointestinal tract of ruminants such as cattle, impose a huge influence on animal performance and physiology. For an example, increased abundance in Firmicutes in the rumen have been linked to increased Average Daily Gain (ADG) and increased feed efficiency. Not only Firmicutes, but a number of other microorganisms living throughout the ruminant gut is well known to have an effect on animal performance parameters such as ADG and average daily feed intake (ADFI) (Myer, 2019)

In the previous two chapters we demonstrated that water restriction can be linked to variations in the composition, and abundance of certain microorganisms in the fecal and rumen microbiomes. With the existing literature suggesting an association between animal performance parameters (such as ADG) and water intake, we conducted the current study to assess the association between population dynamics of the cattle gut microbiome and animal performance under water restricted conditions.

4.2 Materials and Methods

4.2.1 Animal selection for DNA extraction

The % recovery of ADG distributions were calculated (refer section 2.2.3 for further information on % recovery of ADG calculations) separately for each group and the animals that fall in the left tail of the distribution were selected to represent animals with low performance (low percentage recovery of average daily gain – Low PRADG). Those who belonged to the right tail of the % recovery of ADG distribution were selected to represent animals with high performance (high percentage recovery of average daily gain

– High PRADG). A total of 132 animals (66 high performing animals and 66 low performing animals) from the seven groups were selected for fecal and rumen DNA extraction, based on the criteria mentioned above. Rumen and fecal samples collected during water intake and water restriction trials, originating from the selected animals were used for DNA extraction and 16S V4 amplicon sequencing calculated. Refer sections 2.2.4, 2.2.5, 2.2.6 and 2.2.7 for further information on DNA extraction and 16 V4 amplicon sequencing from rumen samples. And refer sections 3.2.3, 3.2.4 and 3.2.5 for further information on DNA extraction and 16 V4 amplicon sequencing from fecal samples.

4.2.2 Data filtering and normalization

Low abundance features, defined as those with less than 10 read counts in at least 20% of the samples were removed prior to data normalization. Rarefying of data based on minimum library size, and Total Sum Scaling was performed as data normalization steps to negate the effect of uneven sequencing depth (Dhariwal et al., 2017). The resulting filtered and normalized data were used for subsequent statistical analyses and data visualizations using MicrobiomeAnalyst.

4.3 Results

4.3.1 Alpha diversity analysis

In order to assess the difference in species richness and evenness between the gut microbiomes of animals with high performance (High PRADG) and animals with low performance (Low PRADG) under water restriction, we conducted an alpha diversity analysis at genus level by grouping the fecal and rumen samples from the water

restriction and baseline trials, based on the performance of the animal they originated from. Alpha diversity measures such as Chao1, Observed ASV, Shannon and Simpsons indices were used for this purpose (Ocejo et al., 2019). Alpha diversity values calculated for baseline and restriction fecal samples are summarized in box plots given in Figure 4.1. and Figure 4.3. The same for rumen baseline and restriction samples are depicted in Figure 4.2 and Figure 4.4.

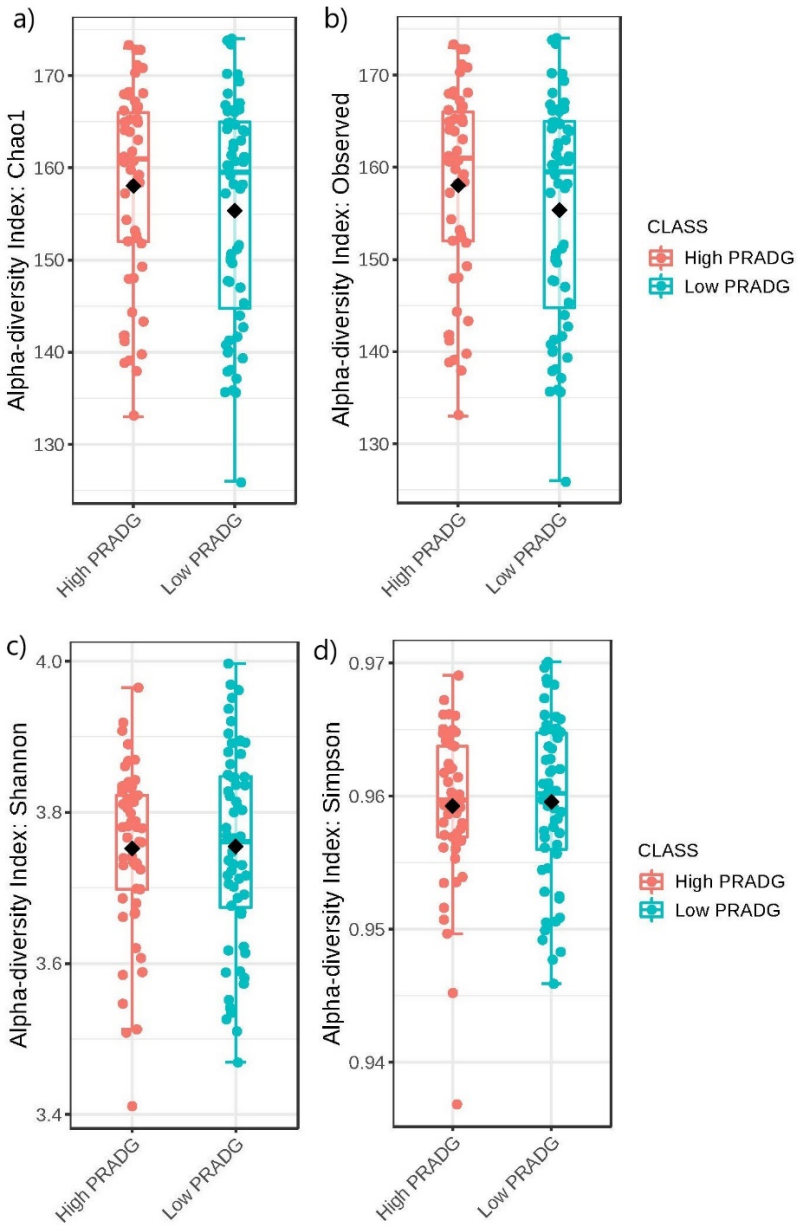


Figure 4.1: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from restriction fecal DNA samples. Fecal samples are grouped into either High PRADG (percentage recovery of average daily gain) or Low PRADG categories and color coded accordingly. Chao1 (a) and Observed ASV (b) indices are based on species richness while Shannon (c) and Simpson (d) indices takes onto account both species richness and evenness.

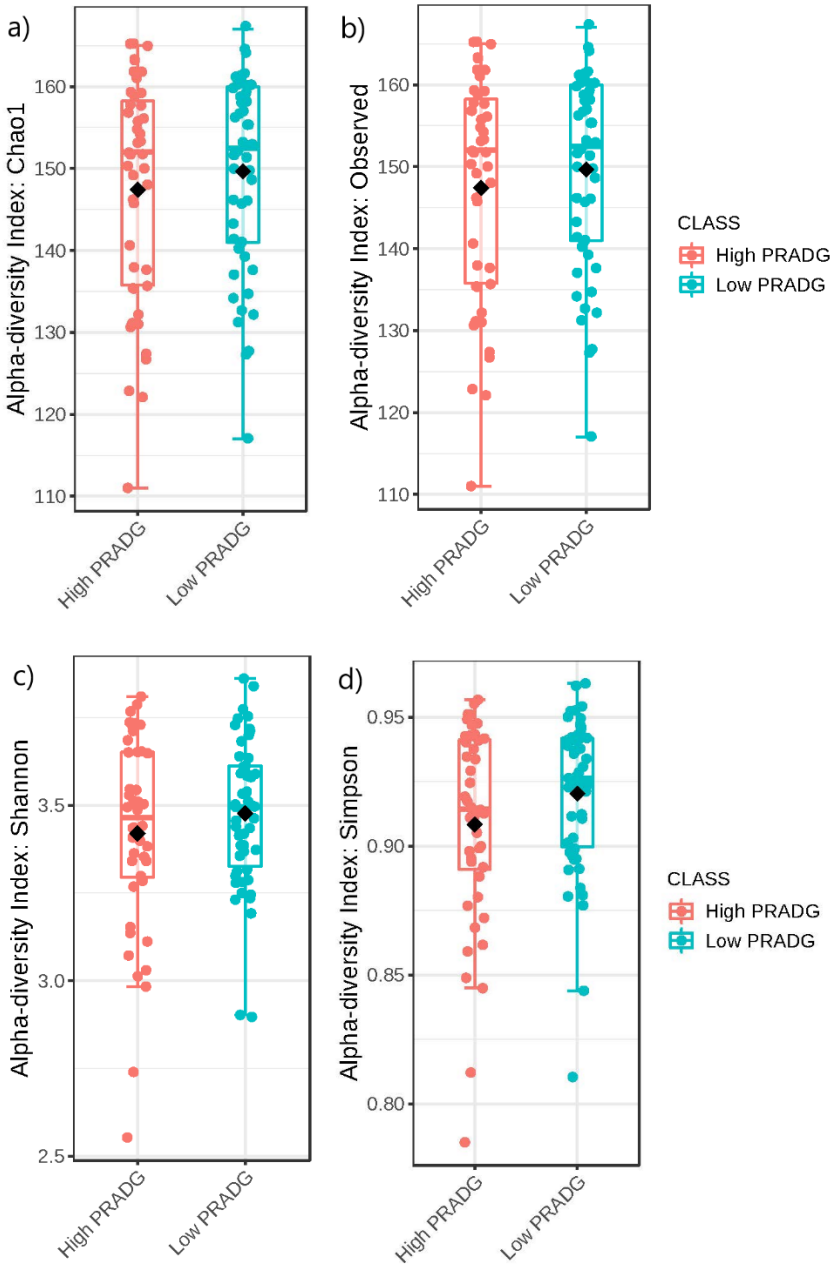


Figure 4.2: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from restriction rumen DNA samples. The rumen samples are grouped into either High PRADG (percentage recovery of average daily gain) or Low PRADG and color coded accordingly. Chao1 (a) and Observed ASV (b) indices are based on species richness while Shannon (c) and Simpson (d) indices takes onto account both species richness and evenness.

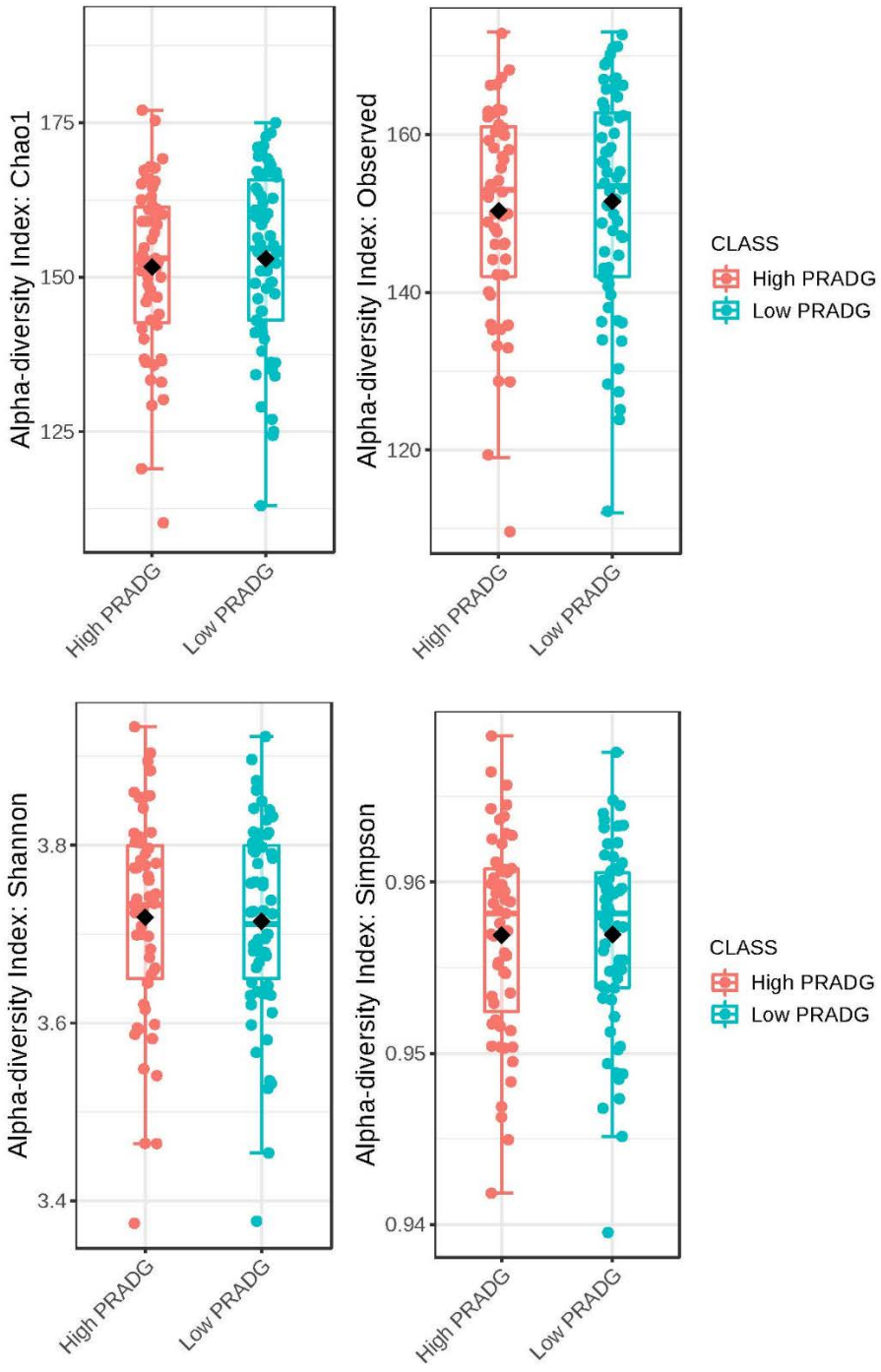


Figure 4.3: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from baseline fecal DNA samples. The rumen samples are grouped into either High PRADG (percentage recovery of average daily gain) or Low

PRADG and color coded accordingly. The Y-axis indicates the respective alpha diversity index for each sample.

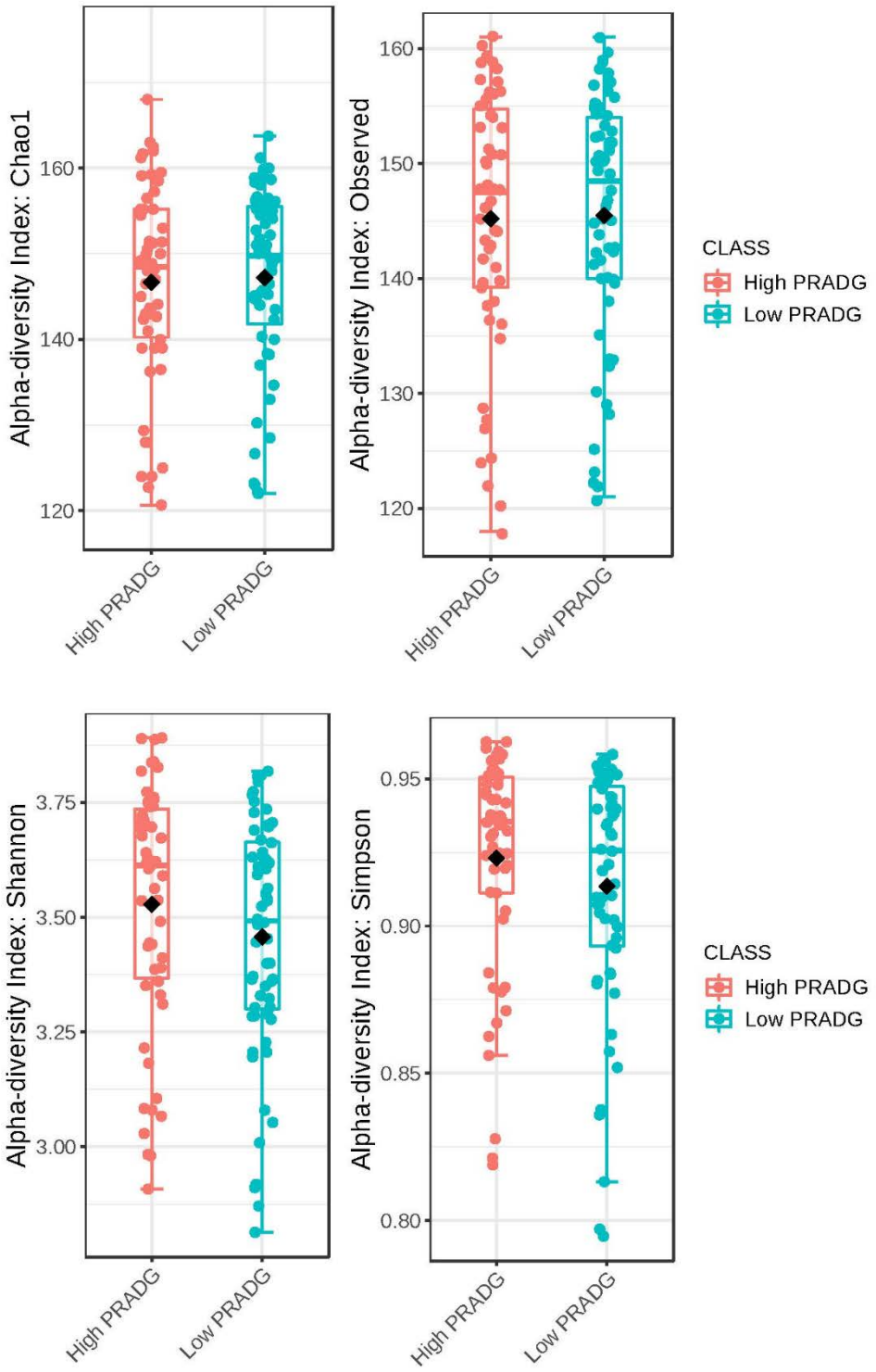


Figure 4.4: Box plots representing the alpha-diversity distributions at genus level for microbial communities recovered from baseline rumen DNA samples. The rumen samples are grouped into either High PRADG (percentage recovery of average daily gain) or Low PRADG and color coded accordingly. The Y-axis indicates the respective alpha diversity index for each sample.

In order to determine the statistical significance of the differences in alpha diversity measures observed between High PRADG and Low PRADG animals, we conducted pairwise comparisons on all alpha diversity indices calculated for both fecal and rumen restriction and baseline samples. Depicted in Table 4.1, Table 4.2, Table 4.3 and Table 4.4 are the results.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	1.0455	0.2980
Chao1	T-test	1.3309	0.1859
Shannon	Man-Whitney	1551	0.8247
Simpson	Man-Whitney	1512	0.6556

Table 4.1: Between group comparisons of alpha diversity indices for restriction fecal samples obtained from High PRADG and Low PRADG animals. T-test was used to measure between group comparisons for the distribution of Observed ASV and Chao1 indices. Man-Whitney test was used for Shannon and Simpson indices.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-0.7926	0.4302
Chao1	T-test	-0.8403	0.4030
Shannon	Man-Whitney	1017	0.5318
Simpson	Man-Whitney	922	0.1786

Table 4.2: Between group comparisons of alpha diversity indices for restriction rumen samples obtained from High PRADG and Low PRADG animals. Between group comparisons for Observed ASV and Chao1 indices were conducted using T-test, while Man-Whitney test was used for between group comparisons of Shannon and Simpson indices.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-0.5102	0.7113
Chao1	T-test	-0.4998	0.6181
Shannon	Man-Whitney	1773	0.7124
Simpson	Man-Whitney	1704	0.9978

Table 4.3: Between group comparisons of alpha diversity indices for baseline fecal samples obtained from High PRADG and Low PRADG animals. T-test was used to measure between group comparisons for the distribution of Observed ASV and Chao1 indices. Man-Whitney test was used for Shannon and Simpson indices

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-0.1343	0.8933
Chao1	T-test	-0.2658	0.7908
Shannon	Man-Whitney	1916	0.0935
Simpson	Man-Whitney	1873	0.1518

Table 4.4: Between group comparisons of alpha diversity indices for baseline rumen samples obtained from High PRADG and Low PRADG animals. T-test was used to measure between group comparisons for the distribution of Observed ASV and Chao1 indices. Man-Whitney test was used for Shannon and Simpson indices

We were not able to observe a significant difference between the alpha diversity of High PRADG and Low PRADG animals for any of the indices used to analyze microbial communities reconstructed from fecal and rumen samples at baseline or restriction water intake.

4.3.2 Beta diversity analysis

In order to compare the species composition of the fecal and rumen microbiomes of low performing and high performing animals under water restriction and baseline conditions, we conducted a beta diversity analysis of the microbial communities reconstructed from rumen and fecal samples obtained during water restriction and water intake (baseline) trials. Bray-Curtis dissimilarity and Jensen-Shannon divergence (Chong et al., 2020) was

used to assess the dissimilarity between the samples from high and low performing animals. The dissimilarity matrices were visualized using NMDS (Non-parametric multidimensional scaling) plots (Luz Calle, 2019) depicted in Figure 4.5, Figure 4.6, Figure 4.7 and Figure 4.8.

The overall clustering pattern was analyzed using permutational analysis of variance (PERMANOVA) in order to make between group comparisons. (Moore et al., 2017). No clear clustering was observed between high performing animals (High PRADG) and low performing animals (Low PRADG) for either rumen or fecal samples. However, for restriction fecal samples there was a significant difference ($p<0.05$) in the overall beta diversity (species composition) between the two groups (Table 4.5). No significant difference between group comparison results were observed for overall beta diversity of restriction rumen samples and baseline rumen/fecal samples from High PRADG and Low PRADG animals (Table 4.6).

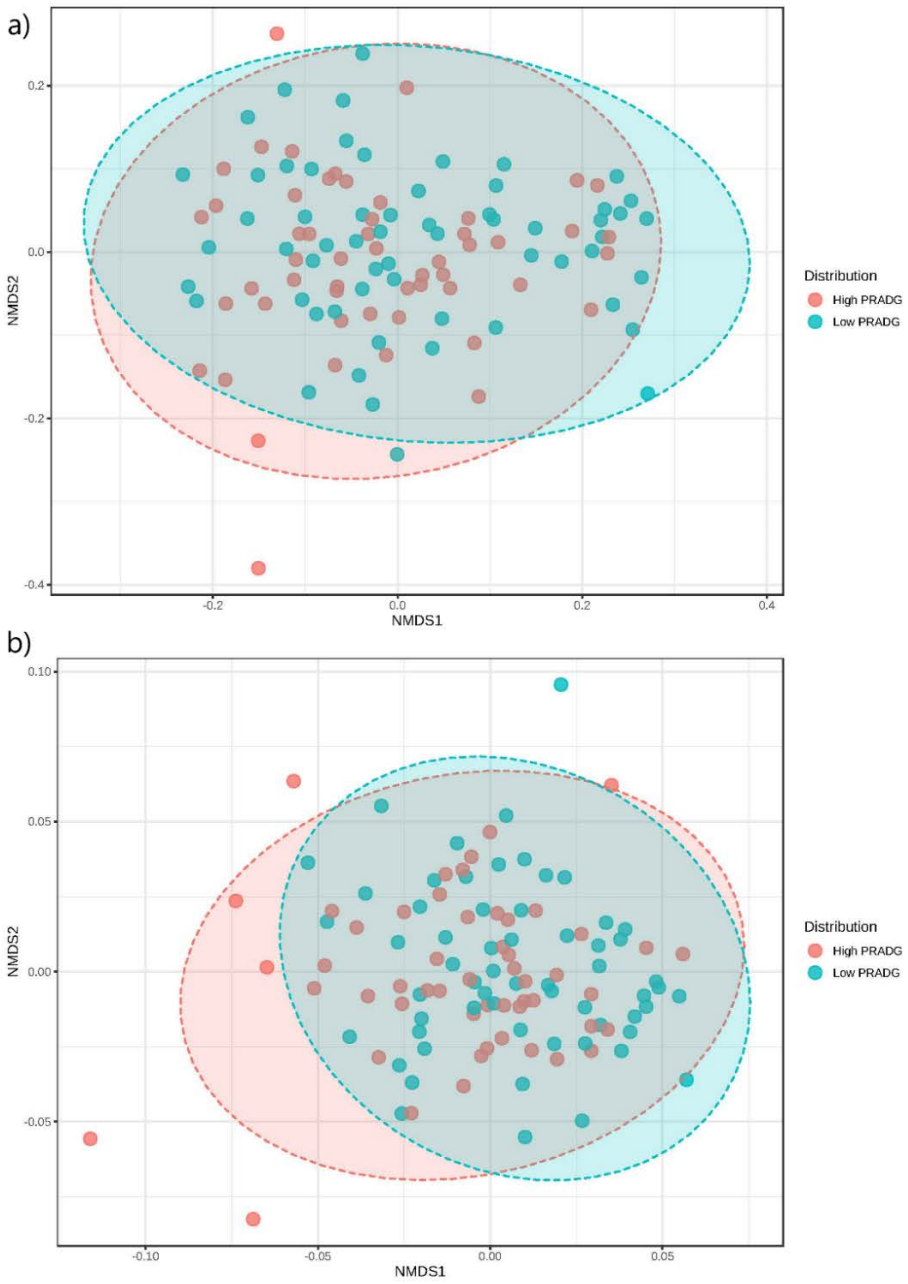


Figure 4.5: NMDS plots illustrating dissimilarity matrices generated using Bray-Curtis (a) and Jensen-Shannon (b) indices in two-dimensional space, for fecal samples obtained from High PRADG and Low PRADG animals. Each blue (Low PRADG animals) and red (High PRADG) point represent the species composition of the entire microbiome present in a given fecal sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups.

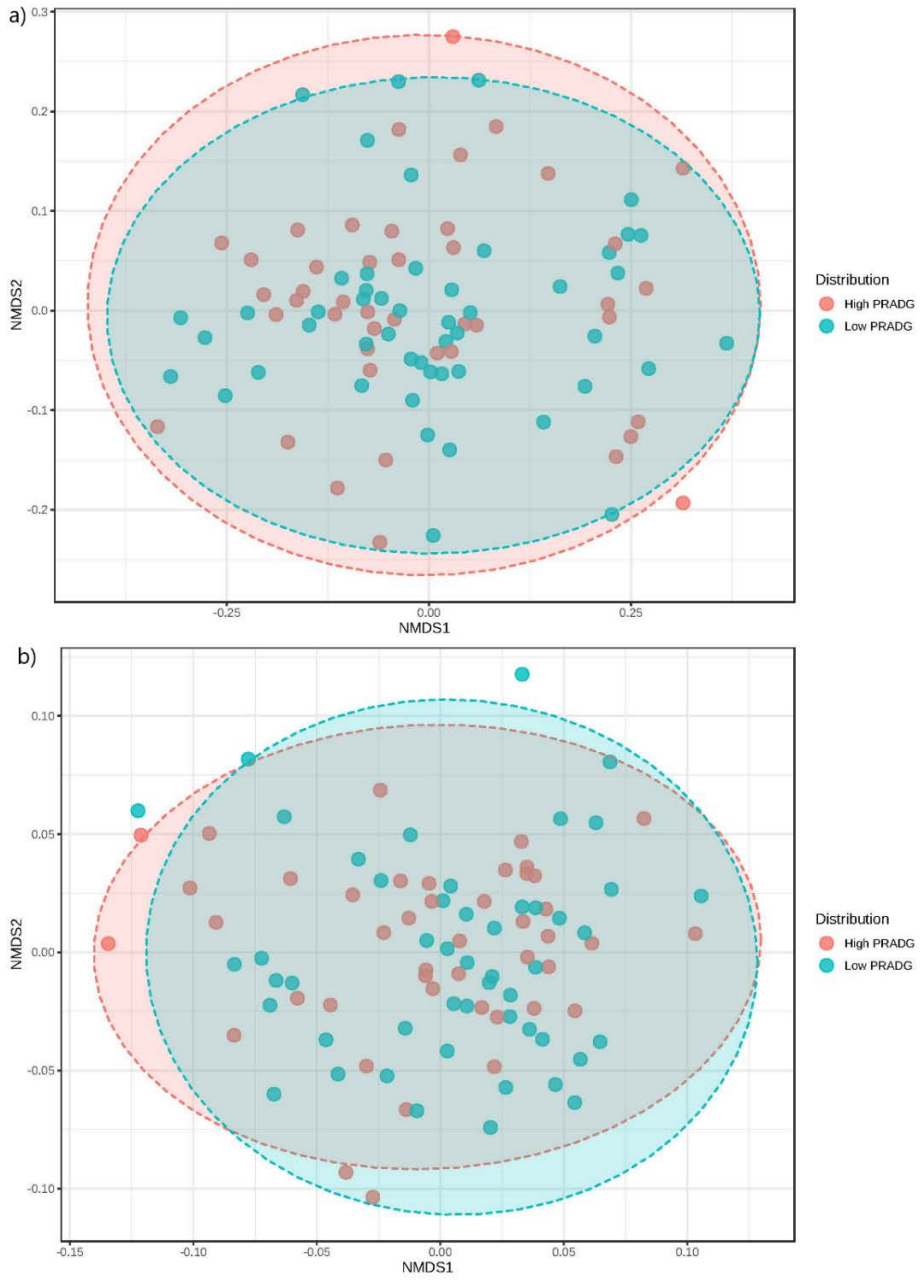


Figure 4.6: NMDS plots illustrating dissimilarity matrices generated using Bray-Curtis (a) and Jensen-Shannon (b) indices in two-dimensional space, for rumen samples obtained from High PRADG and Low PRADG animals. Each blue (Low PRADG animals) and red (High PRADG) point represent the species composition of the entire microbiome present in a given fecal sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups.

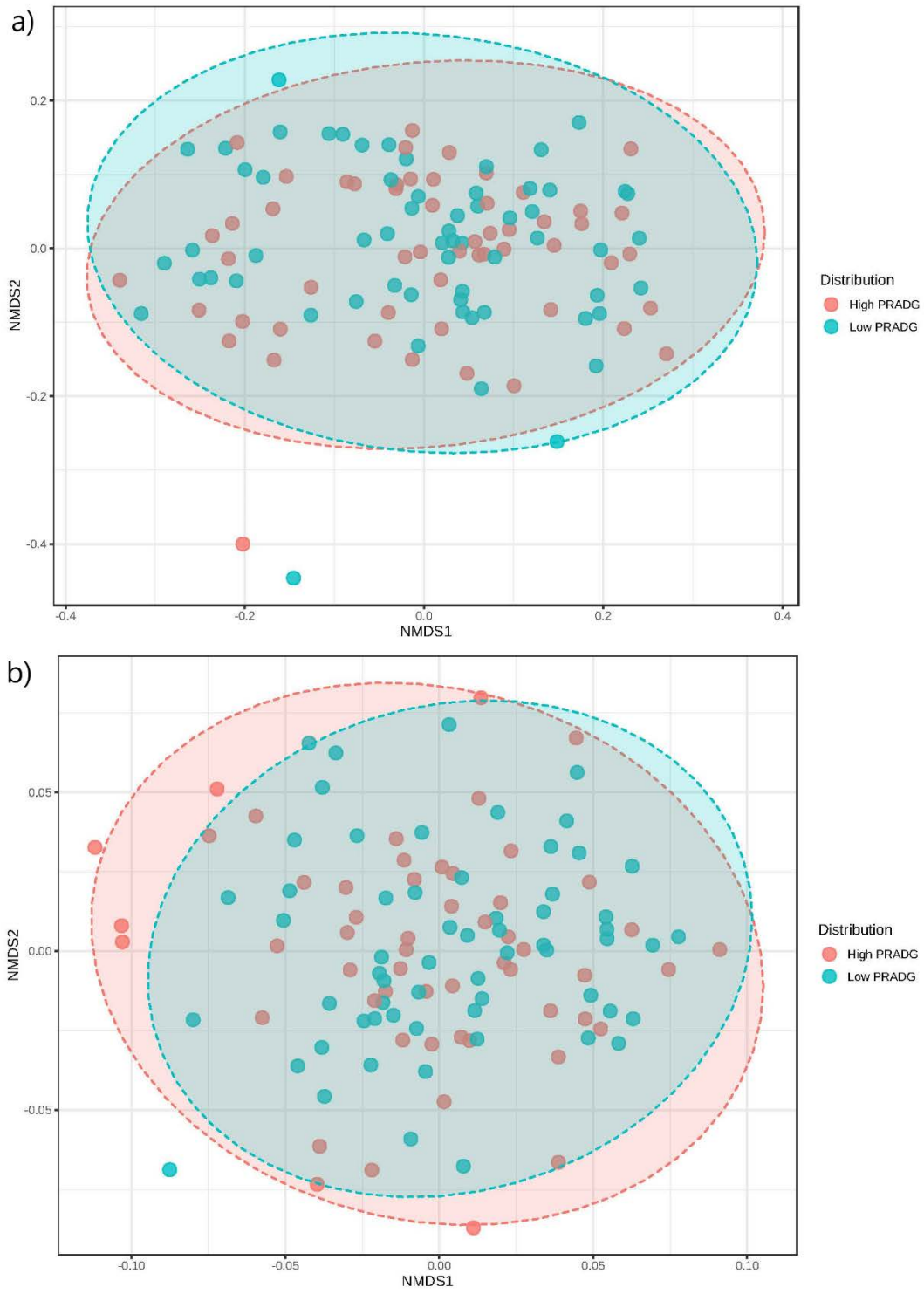


Figure 4.7: NMDS plots illustrating dissimilarity matrices generated using Bray-Curtis (a) and Jensen-Shannon (b) indices in two-dimensional space, for baseline fecal samples

obtained from High PRADG and Low PRADG animals. Each blue (Low PRADG animals) and red (High PRADG) point represent the species composition of the entire microbiome present in a given fecal sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups.

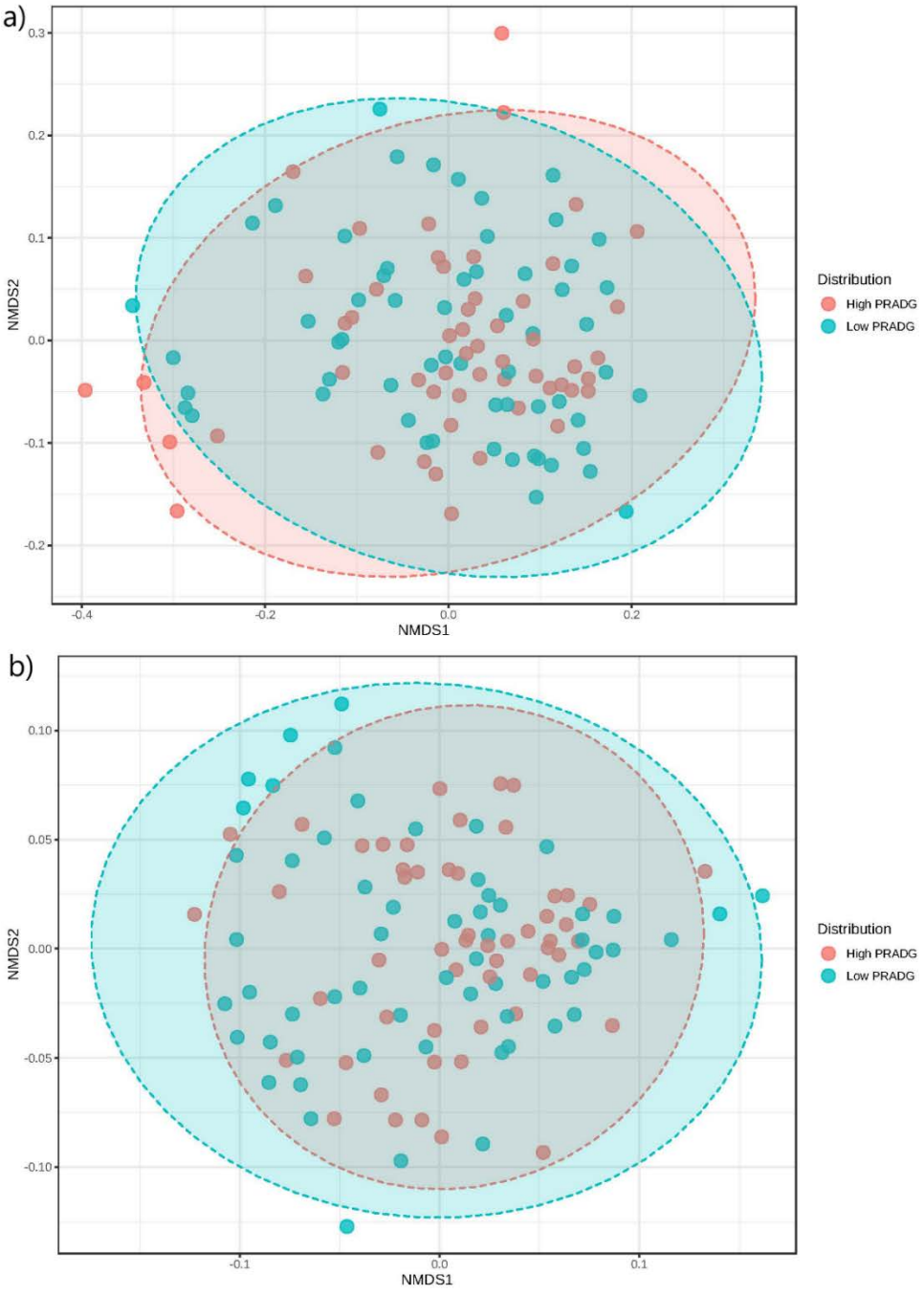


Figure 4.8: NMDS plots illustrating dissimilarity matrices generated using Bray-Curtis (a) and Jensen-Shannon (b) indices in two-dimensional space, for baseline rumen samples obtained from High PRADG and Low PRADG animals. Each blue (Low PRADG animals) and red (High PRADG) point represent the species composition of the entire microbiome present in a given fecal sample. The ellipses mark the 95% confidence interval surrounding the centroid of each of the two groups.

Sample Type	Distance method	F-value	R-squared	p-value
Restriction fecal samples	Bray-Curtis dissimilarity	2.4333	0.0214	0.011
	Jensen-Shannon diversgence	4.1053	0.0356	0.004
Restriction rumen samples	Bray-Curtis dissimilarity	1.4488	0.0155	0.168
	Jensen-Shannon diversgence	1.2333	0.0132	0.308

Table 4.5: The results of the PERMANOVA test for between group comparison of High PRADG and Low PRADG groups based on the microbial community composition of restriction rumen and fecal samples. For restriction fecal samples the community composition differs significantly ($p<0.05$) for both the distance methods used. For restriction rumen samples, no significant difference in between group comparisons was observed ($p>0.05$), for either of the distance methods used.

Sample Type	Distance method	F-value	R-squared	p-value
Baseline fecal samples	Bray-Curtis dissimilarity	0.6887	0.0059	0.737
	Jensen-Shannon divergence	0.3882	0.0033	0.767
Baseline rumen samples	Bray-Curtis dissimilarity	1.2428	0.0109	0.244
	Jensen-Shannon divergence	0.4762	0.0042	0.638

Table 4.6: Between group comparisons of distance methods used to measure beta diversity of baseline rumen and fecal microbiome of low performing and high performing

animals. None of the comparisons between high and low performing animals revealed a significant difference in their beta diversity.

4.3.3 Microbial community composition

The taxonomic composition of the microbial communities recovered from fecal and rumen samples from the water restriction trial were visualized using relative abundance graphs. The samples were grouped based on animal performance (i.e. High PRADG animals and Low PRADG animals) and the data were summarized at genus level.

As depicted in Figure 4.8, The taxonomy of the 10 most abundant genera remained the same between High PRADG and Low PRADG animals for both rumen and fecal samples obtained during the water restriction trial. However, the relative abundance of some of the genera differed significantly between the two groups. For fecal samples, we observed a significant decrease ($p < 0.05$) in the abundance of *Christensenellaceae_R_7_group* and *Paeniclostridium* in High PRADG animals (compared to Low PRADG animals), while a significant increase ($p<0.05$) in abundance was observed for *Methanobrevibacter* (Table 4.7). For rumen restriction samples, *Christensenellaceae_R_7_group* and *Ruminococcus* showed a significant decrease ($p<0.05$) in abundance in High PRADG animals, whereas the relative abundance of *Prevotella* showed a significant ($p<0.05$) increase (Table 4.8).

The baseline fecal microbiome did not show a significant difference in terms of overall species composition as well as the composition and the abundance of the most predominant genera observed between high performing and low performing animals

(Table 4.9). Only the abundance of *Christensenellaceae_R_7_group* and *Rikenellaceae_RC9_gut_group* differed significantly between the baseline rumen microbiomes of high performing and low performing animals (Table 4.10).

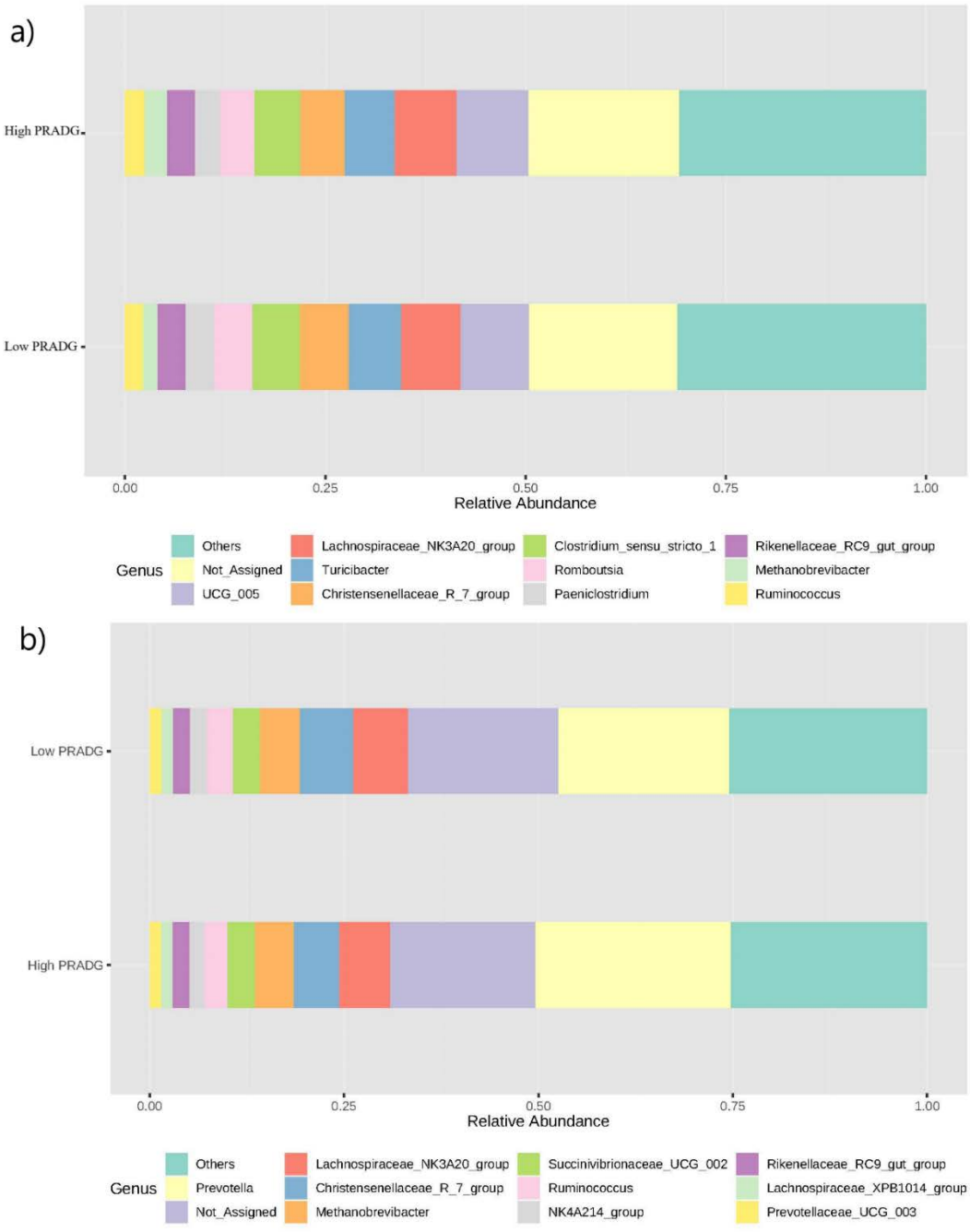


Figure 4.8: Relative abundance bar graphs depicting the 10 most abundant genera for fecal (a) and rumen (b) samples from the water restriction trial. The results are grouped according to animal performance, i.e High PRADG and Low PRADG animals.

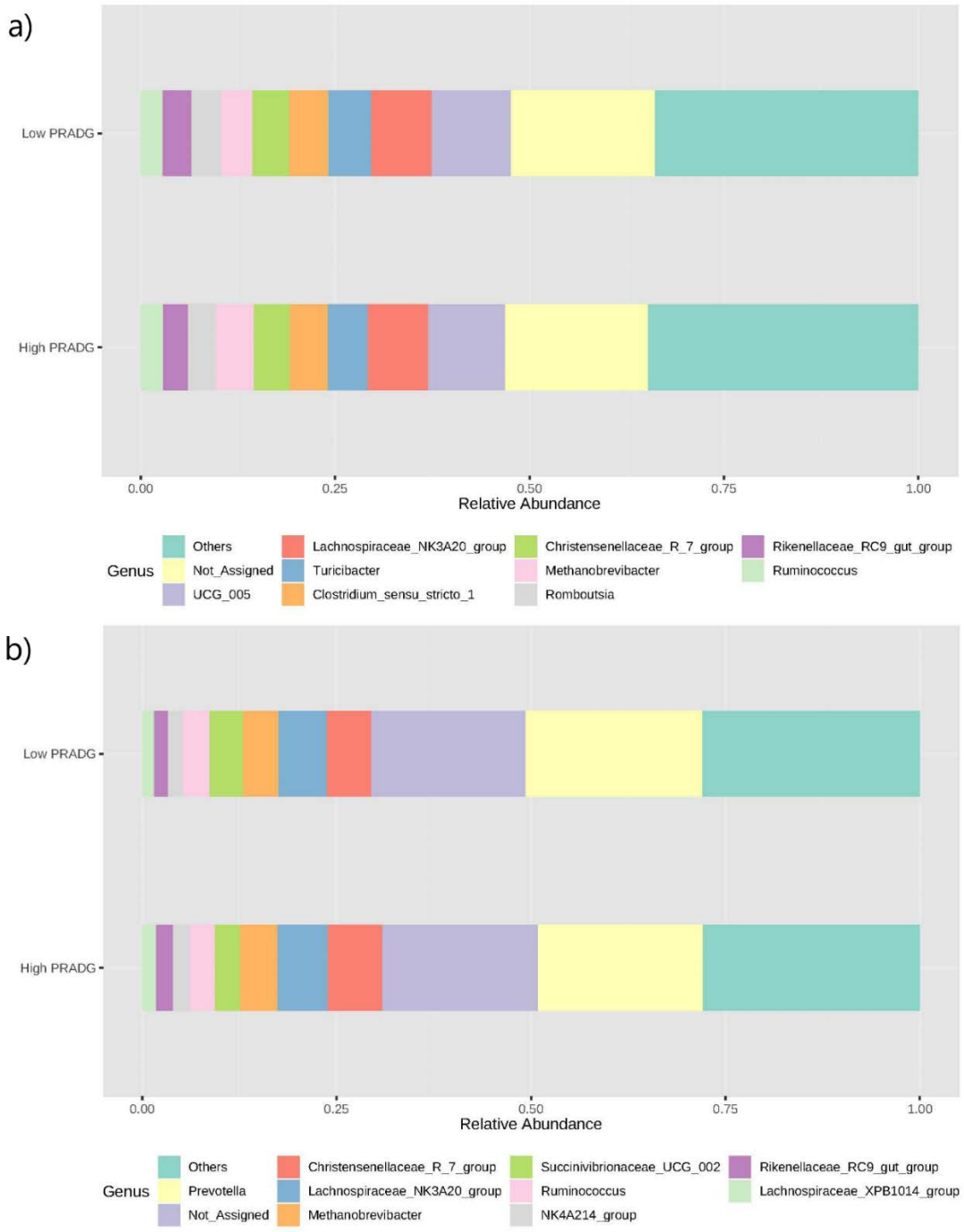


Figure 4.9: Relative abundance bar graphs depicting the 10 most abundant genera for fecal (a) and rumen (b) samples from the water intake (baseline) trial. The results are grouped according to animal performance, i.e High PRADG and Low PRADG animals.

Genus	Low PRADG	High PRADG	P-value
UCG_005	8.47%	8.92%	0.358
Lachnospiraceae_NK3A20_group	7.48%	7.80%	0.490
Turicibacter	6.48%	6.15%	0.395
Christensenellaceae_R_7_group	6.17%	5.61%	0.010*
Clostridium_sensu_stricto_1	5.95%	5.70%	0.368
Romboutsia	4.67%	4.23%	0.072
Paeniclostridium	3.67%	3.20%	0.047*
Rikenellaceae_RC9_gut_group	3.42%	3.41%	0.850
Methanobrevibacter	1.87%	2.88%	0.003*
Ruminococcus	2.22%	2.43%	0.218

(*) – $p < 0.05$

Table 4.7: Relative abundances of the 10 most abundant genera in fecal samples obtained during the water restriction trial summarized according to animal performance (High PRADG and Low PRADG).

Genus	High PRADG	Low PRADG	P-value
Prevotella	25.04%	21.92%	0.040*
Lachnospiraceae_NK3A20_group	6.56%	7.06%	0.315
Christensenellaceae_R_7_group	5.90%	6.91%	0.037*
Methanobrevibacter	5.00%	5.13%	0.781
Succinivibrionaceae_UCG_002	3.50%	3.41%	0.926
Ruminococcus	2.94%	3.35%	0.030*
NK4A214_group	2.00%	2.28%	0.079
Rikenellaceae_RC9_gut_group	2.09%	2.10%	0.984
Lachnospiraceae_XPB1014_group	1.54%	1.53%	0.963
Prevotellaceae_UCG_003	1.45%	1.50%	0.640

(*) – $p < 0.05$

Table 4.8: Relative abundances of the 10 most abundant genera in rumen samples obtained during water restriction trial summarized according to animal performance (High PRADG and Low PRADG).

Genera	High PRADG	Low PRADG	P-values
UCG_005	9.88%	10.09%	0.641
Lachnospiraceae_NK3A20_group	7.77%	7.87%	0.791
Turicibacter	5.08%	5.44%	0.465
Clostridium_sensu_stricto_1	4.88%	5.02%	0.752
Christensenellaceae_R_7_group	4.73%	4.80%	0.705
Methanobrevibacter	4.88%	3.97%	0.096
Romboutsia	3.54%	3.82%	0.336
Rikenellaceae_RC9_gut_group	3.27%	3.79%	0.077
Ruminococcus	2.84%	2.77%	0.794
Paeniclostridium	2.26%	2.37%	0.651

Table 4. 9: Relative abundances of the 10 most abundant genera in fecal samples obtained during the water intake trial summarized according to animal performance (High PRADG and Low PRADG).

Genera	High PRADG	Low PRADG	P-value
Prevotella	21.12%	22.66%	0.3644
Christensenellaceae_R_7_group	7.07%	5.78%	0.0054*
Lachnospiraceae_NK3A20_group	6.52%	6.25%	0.5205
Methanobrevibacter	4.76%	4.50%	0.6217
Succinivibrionaceae_UCG_002	3.18%	4.24%	0.2088
Ruminococcus	3.22%	3.43%	0.3136
NK4A214_group	2.21%	2.00%	0.1810
Rikenellaceae_RC9_gut_group	2.16%	1.70%	0.0015*
Lachnospiraceae_XPB1014_group	1.80%	1.58%	0.2854
Prevotellaceae_UCG_003	1.58%	1.61%	0.9351

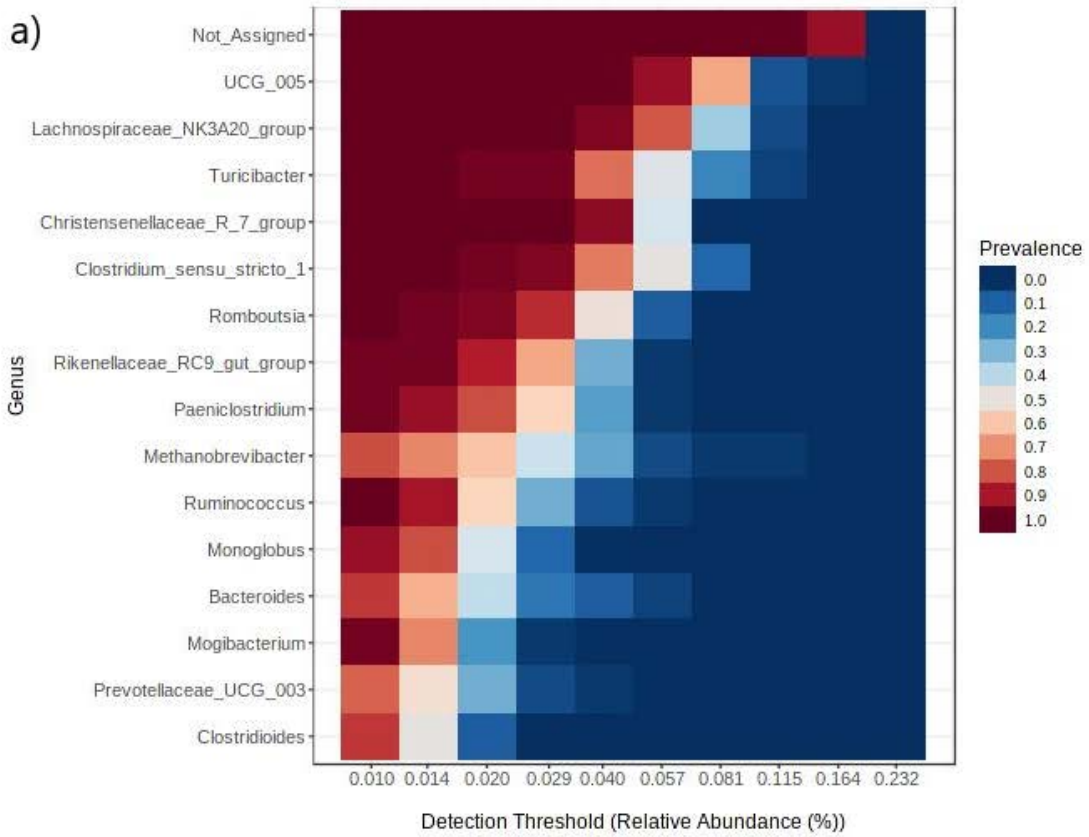
(*) $p<0.05$

Table 4.10: Relative abundances of the 10 most abundant genera in rumen samples obtained during the water intake trial summarized according to animal performance (High PRADG and Low PRADG).

4.3.4 The core-microbiome

In order to determine whether there is a difference in the most prevalent microbial taxa in the gut microbiome of animals who perform well during water restriction, we derived a core microbiome for fecal and rumen samples obtained during the water restriction trial. The core microbiome was defined as the taxa that are present in at least 70% of the samples considered, at a threshold relative abundance of 0.01% or above.

As depicted in figures 4.10, 4.11, 4.12 and 4.13, for both rumen and fecal samples, the core microbiome for High PRADG and Low PRADG animals essentially looked the same with two exceptions in the restriction fecal microbiome and baseline fecal microbiome. M*ethanobrevibacter* present in the fecal core microbiome of High PRADG animals, was replaced by *Familly_XIII_AD3011_group* in the Low PRADG animals. Prevotellaceae_UCG_003 appeared additionally in the baseline fecal microbiome of low performing animals compared to the same in high performing animals.


a)
Not_Assigned
UCG_005
Lachnospiraceae_NK3A20_group
Turicibacter
Christensenellaceae_R_7_group
Clostridium_sensu_stricto_1
Romboutsia
Rikenellaceae_RC9_gut_group
Paeniclostridium
Methanobrevibacter
Ruminococcus
Monoglobus
Bacteroides
Mogibacterium
Prevotellaceae_UCG_003
Clostridioides
Genus
0.010 0.014 0.020 0.029 0.040 0.057 0.081 0.115 0.164 0.232
Detection Threshold (Relative Abundance (%))
Prevalence
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

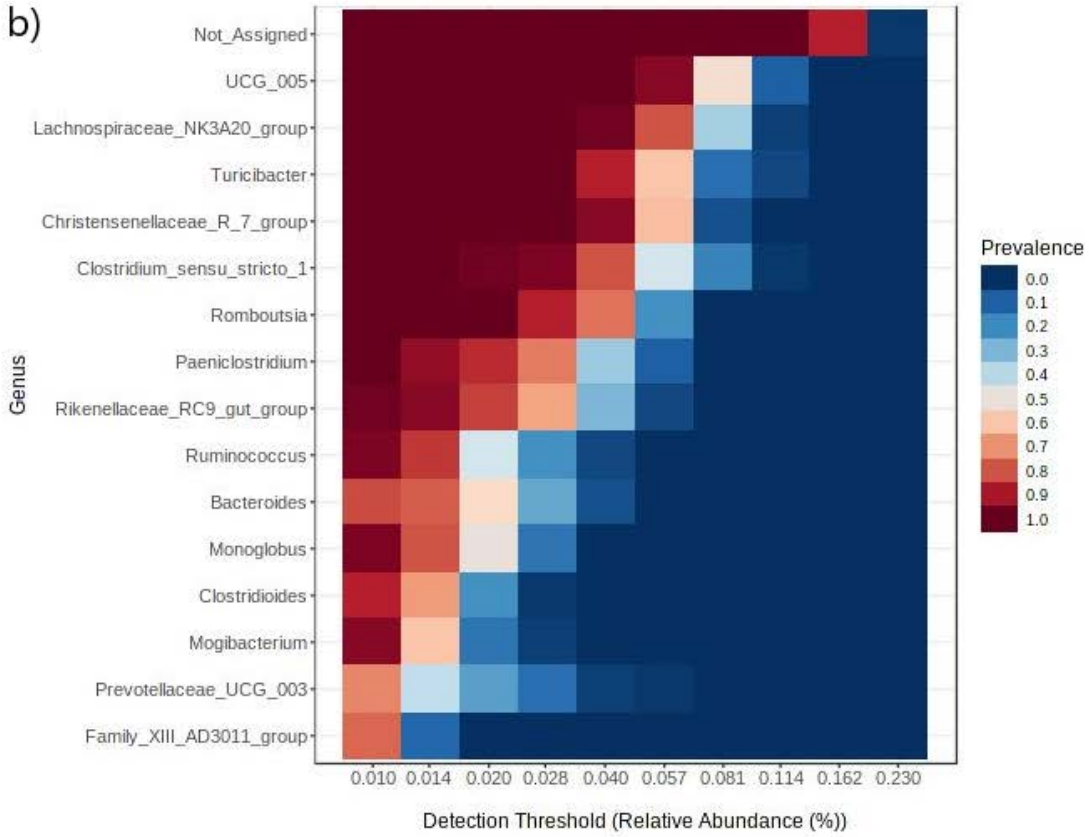

b)
Not_Assigned
UCG_005
Lachnospiraceae_NK3A20_group
Turicibacter
Christensenellaceae_R_7_group
Clostridium_sensu_stricto_1
Romboutsia
Paeniclostridium
Rikenellaceae_RC9_gut_group
Ruminococcus
Bacteroides
Monoglobus
Clostridioides
Mogibacterium
Prevotellaceae_UCG_003
Family_XIII_AD3011_group
Genus
0.010 0.014 0.020 0.028 0.040 0.057 0.081 0.114 0.162 0.230
Detection Threshold (Relative Abundance (%))
Prevalence
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Figure 4.10: The heat map illustrating the relative abundance and prevalence of the microbial taxa defining the core-microbiomes recovered from fecal samples obtained during the water restriction trial from High PRADG (a) and Low PRADG (b) animals. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria for microbial taxa to be considered as a part of the core microbiome.

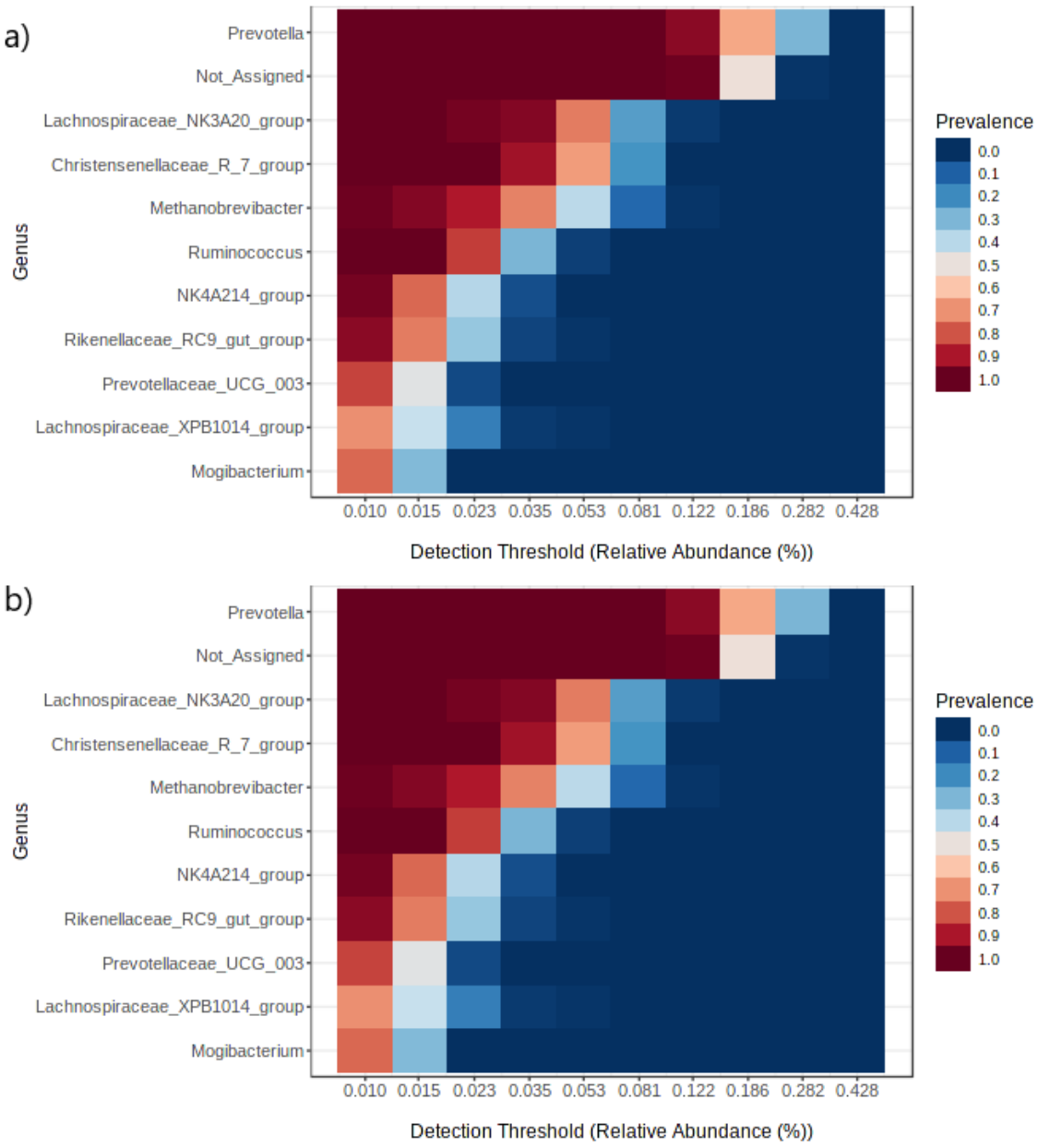


Figure 4.11: Heat maps illustrating the relative abundance and prevalence of the microbial taxa defining the core-microbiomes recovered from restriction rumen samples of High PRADG (a) and Low PRADG (b) animals. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria for microbial taxa to be considered as a part of the core microbiome.

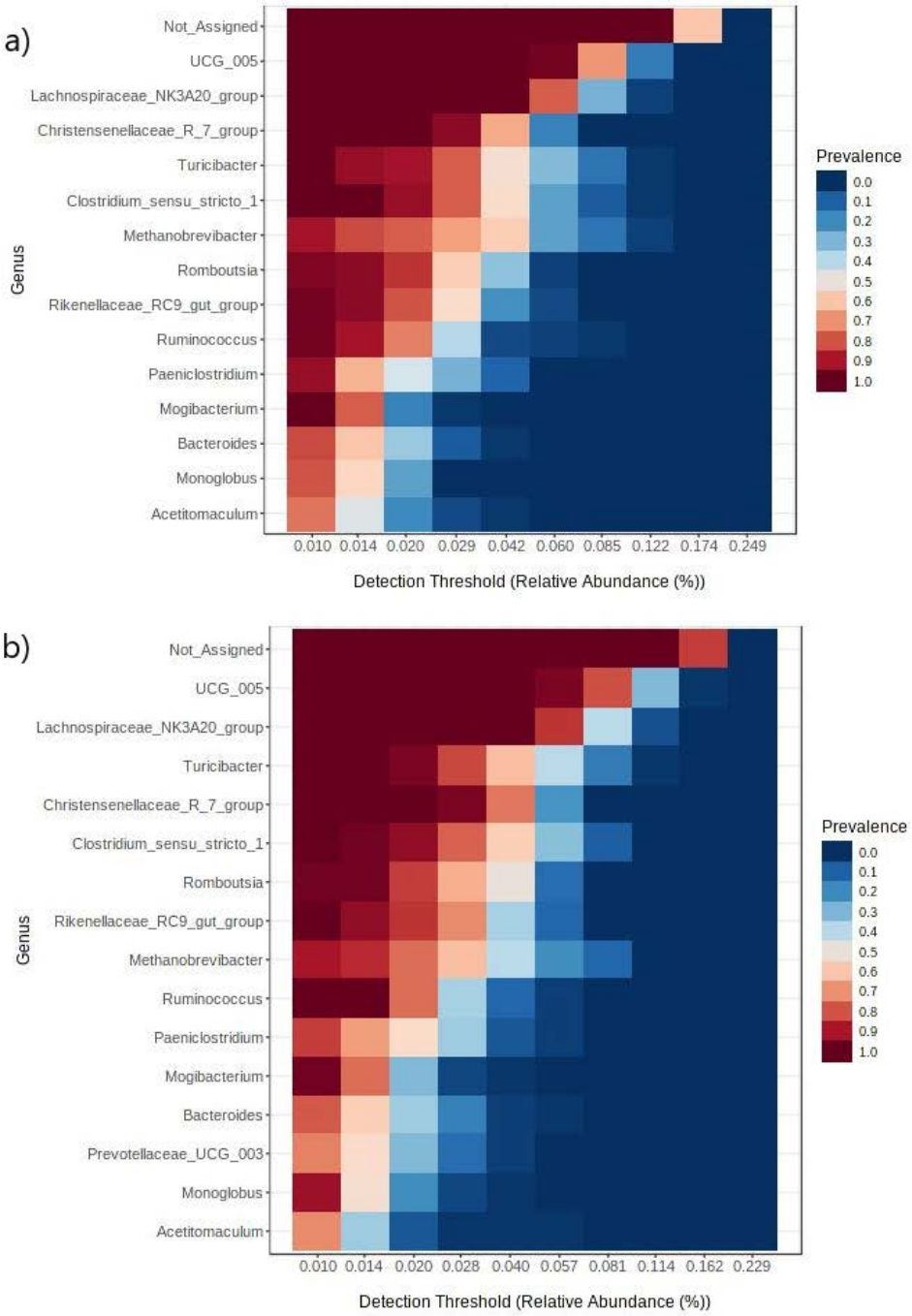


Figure 4.12: Heat maps illustrating the relative abundance and prevalence of the microbial taxa defining the core-microbiomes recovered from baseline fecal samples of High PRADG (a) and Low PRADG (b) animals. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria for microbial taxa to be considered as a part of the core microbiome.

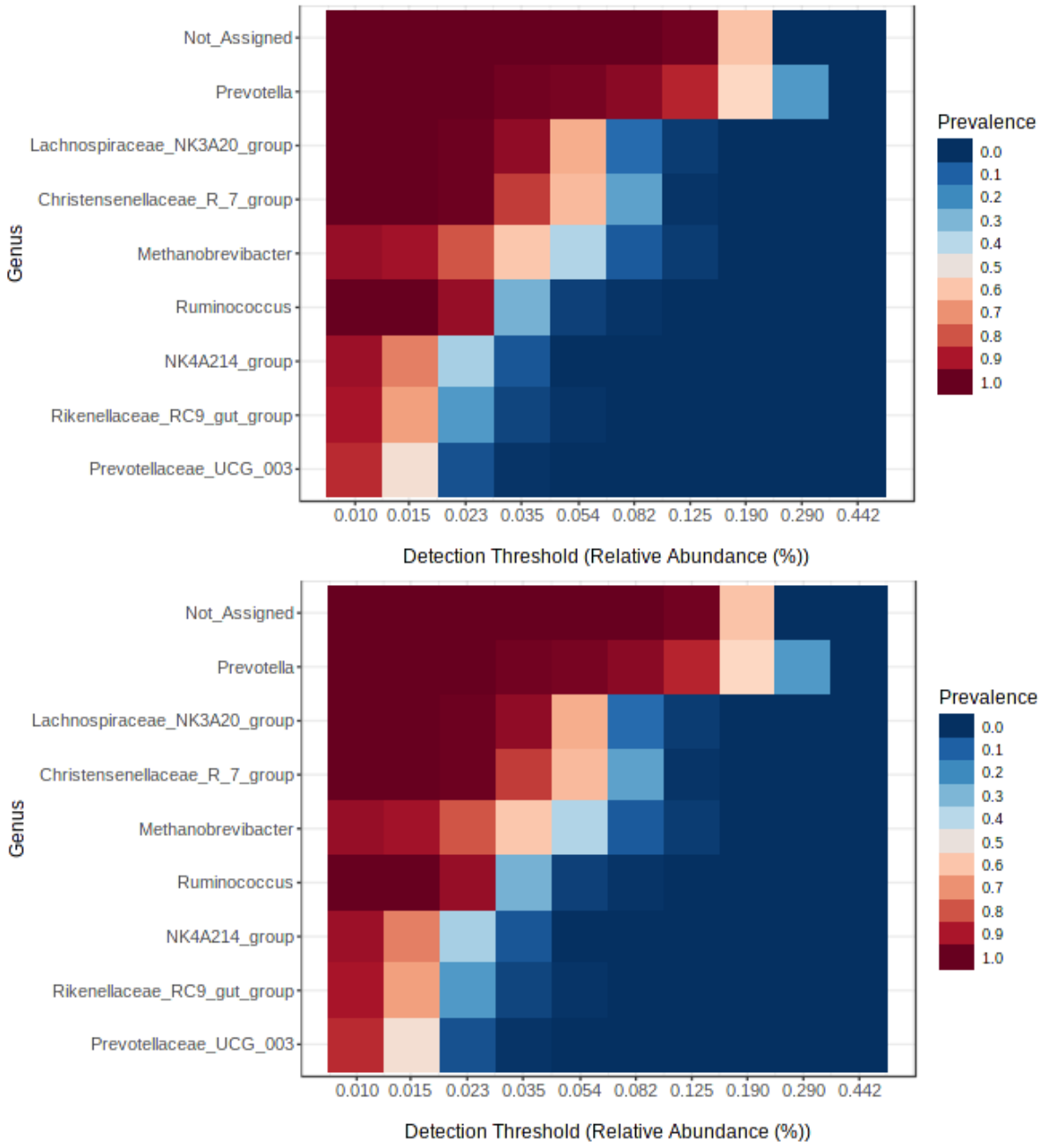


Figure 4.13: Heat maps illustrating the relative abundance and prevalence of the microbial taxa defining the core-microbiomes recovered from baseline rumen samples of High PRADG (a) and Low PRADG (b) animals. A threshold of at least 70% prevalence and 0.01% relative abundance was used as the selection criteria for microbial taxa to be considered as a part of the core microbiome.

4.3.5 Biomarker analysis based on animal performance.

In order to identify microbial biomarkers (differentially abundant taxa with biological significance) that are enriched in high performing (High PRADG) and low performing (Low PRADG) animals, a LEfSe (Linear Discriminant Analysis Effect Size) analysis (Segata et al., 2011) was conducted using fecal and rumen samples from the water restriction trial. For restriction fecal samples, we were able to identify 23 differentially abundant taxa between the two animal groups, and only 5 differentially abundant taxa based on restriction rumen samples. A P-value<0.05 and an LDA score> 2.0 was used as the selection criteria for biomarkers.

As depicted in Figure 4.14a, out of the 15 most enriched biomarkers identified by LEfSe based on the restriction fecal samples, 7 were differentially abundant in High PRADG animals, namely*; Lachnospiraceae_UCG_002, Blautia, UCG_001, Mehthanosphaera, UCG_001, Methanosphaera, Coprococcus, Acetitomaculum,* and *Methanobrevibacter*. The remaining eight were differentially abundant in Low PRADG animals, namely; *Christensenellaceae_R_7_group, Paenoclostridum, Rombutsia, Clostridoided, UCG_002, Akkermansia, Lactobacillus,* and *dgA_11_gut_group.* For restriction rumen samples, only five genera immerged as differentially abundant taxa, namely; *Saccharofermentans, Mogibacterium, Lachnoclostridium, Lachnospiraceae_FCS020_group,* and *Prevotella*. *Saccharofermentans, Mogibacterium,* and *Lachnoclostridium* were associated with the rumen microbiome of Low PRADG animals, while *Lachnospiraceae_FCS020_group,* and *Prevotella* were associated with the rumen microbiome of High PRADG animals.

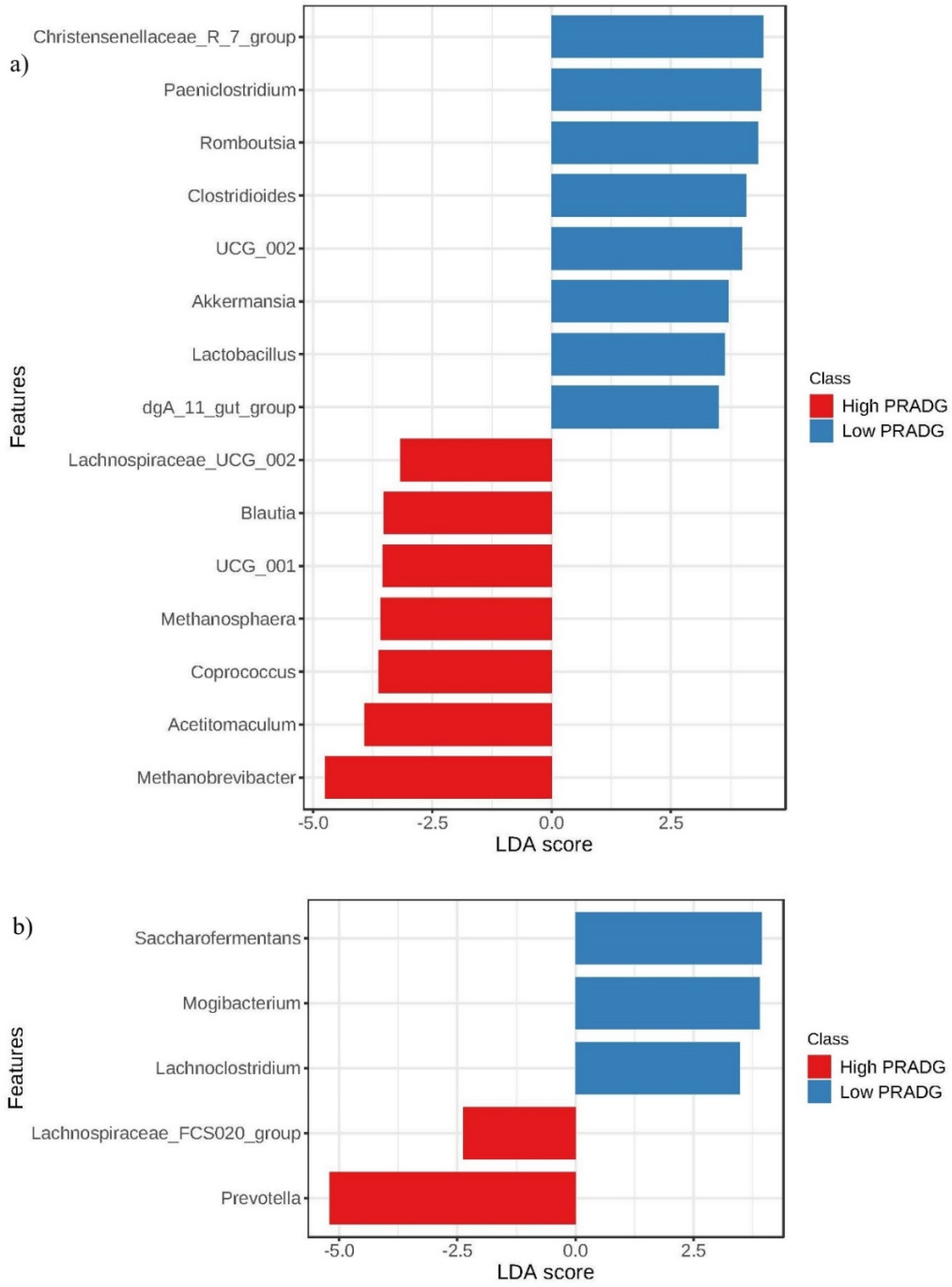


Figure 4.14: Differentially abundant bacterial genera in the gut microbiome of High PRADG and Low PRADG animals during water restriction. Plot **(a)** depicts the 15 most enriched biomarkers (out of the 23 total

biomarkers identified by LEfSe analysis) associated with the fecal microbiome of High PRADG animals and Low PRADG animals during water restriction. Plot (b) depicts the biomarkers associated with High PRADG animals and Low PRADG animals based on their rumen microbiome during water restriction. The selection criteria for differentially abundant features were a P-value < 0.05 and LDA score > 2.0. The blue bars indicate LDA score of biomarkers associated with Low PRADG animals and the red bars indicate LDA scores of biomarkers associated with High PRADG animals. (Segata et al., 2011; Chong et al., 2020).

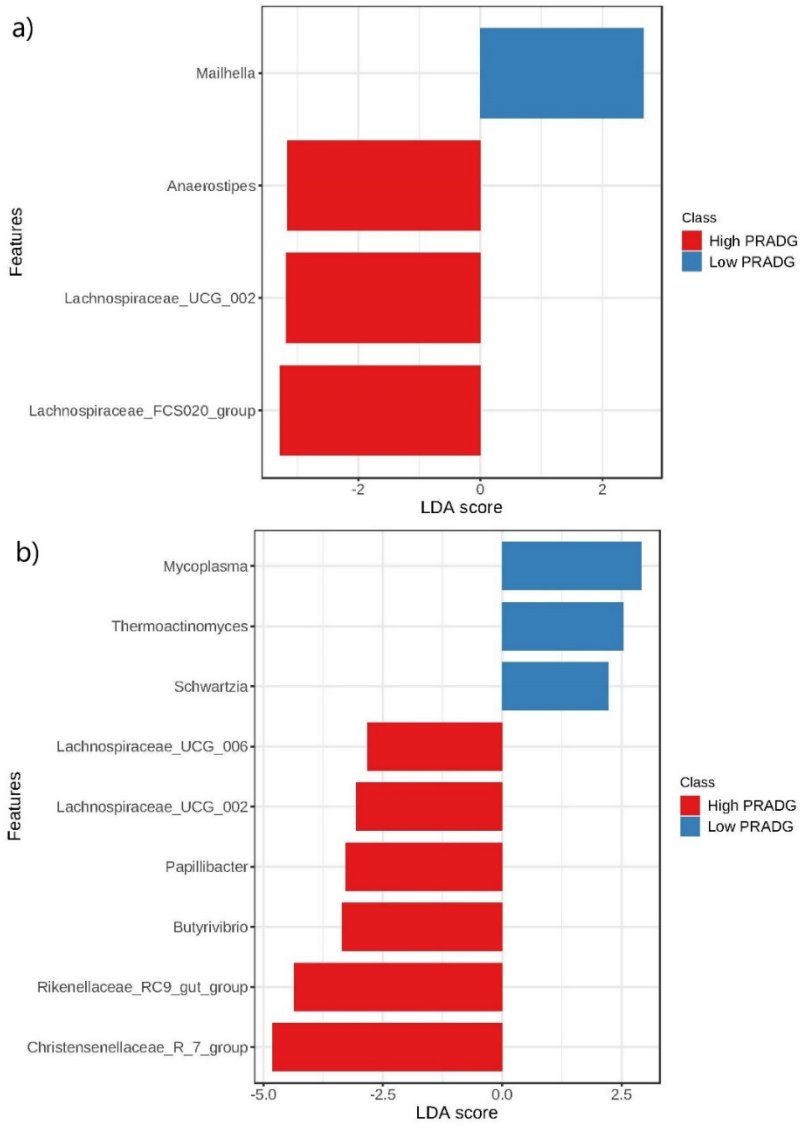


Figure 4.15: Differentially abundant bacterial genera in the gut microbiome of High PRADG and Low PRADG animals. Plot **(a)** depicts enriched biomarkers associated with the fecal microbiome of High PRADG animals and Low PRADG animals during baseline water intake. Plot (b) depicts the biomarkers associated with High PRADG animals and Low PRADG animals based on their rumen microbiome during baseline water intake. The selection criteria for differentially abundant features were a P-value < 0.05 and LDA score > 2.0. The blue bars indicate LDA score of biomarkers associated with Low

PRADG animals and the red bars indicate LDA scores of biomarkers associated with High PRADG animals. (Segata et al., 2011; Chong et al., 2020).

LEfSe analysis on the baseline fecal microbiome revealed four biomarkers with only *Mailhela* associated with low performing animals, while *Anerostipes, Lachnospiraceae_UCG_002 and Lachnospiraceae_FCS_020_group* were associated with high performing animals (Figure 4.15(a)). The baseline rumen microbiome revealed 9 biomarkers associated with high and low performing animals as depicted in (Figure 4.15(b)).

Random forest (RF) analysis (Breiman, 2001) implemented in MicrobiomeAnalyst (Chong et al., 2020) was used to further elucidate the differentially abundant taxa between Low PRADG and High PRADG animals using fecal and rumen samples from the water restriction trial, at genus level. By running RF analysis on restriction fecal samples, we were able to identify *Parvibacter, Streptococcus, Lacotbacillus, Methanobrevibactor, Blautia, Coprococcus,* and *Lachnospiraceae-UCG-002* as differentially abundant species associated with High PRADG animals. *While p_1088_a5_gut_group, Akkermansia, Christensenellaceae_R7_group, UCG-002, Denitrobacterium, Geobacillus, Caproicipruducens,* and *Clostridioides* were associated with Low PRADG animals (Figure 4.9).

For restriction rumen samples, only *Provetollaceae_Ga6A1_group* and *Aeriscardovia* out of the top 15 differentially abundant taxa, were identified as features associated with High PRADG animals. The rest of the top 15 differentially abundant taxa identified,

namely; *Saccharofermentans, Lachnospiraceae_NKA416_group, Oscillospira, Christensenellaceae_R7_group, Lactobacillus, Flexilinea, Veillonellacea_UCG-001, Lachnospiraceae_FE2018_group, Sacharopolyspora, Prevetollaceae_UCG-003, Elusimicrobium, Syntrophococcus*, and *Papillibacter* were all associated with Low PRADG animals (Figure 4.10). Random Forest analysis on baseline and rumen and fecal microbiome was not able to identify significant biomarkers.

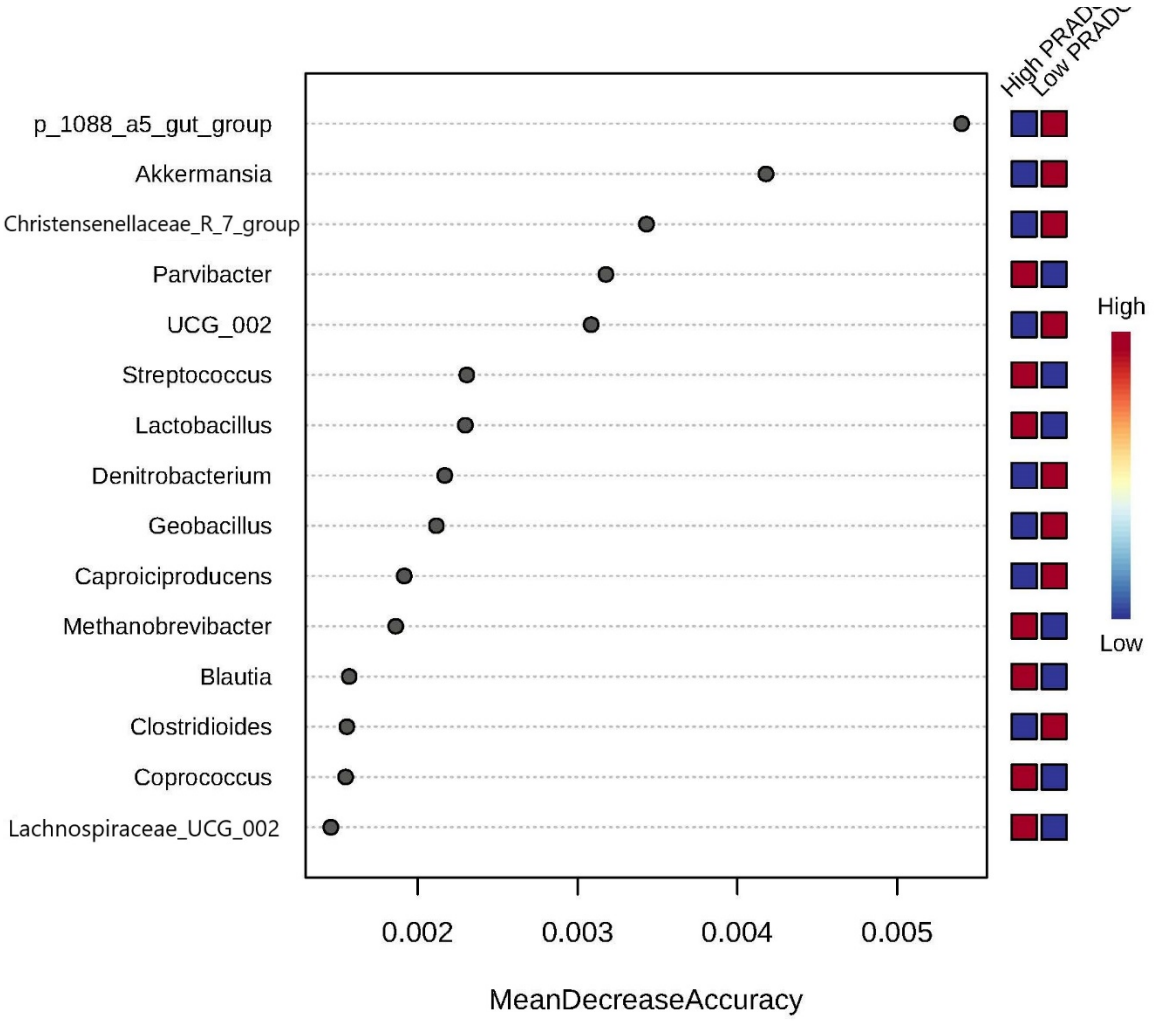


Figure 4.16: Differentially abundant taxa between High PRADG and Low PRADG animals in the restriction fecal microbiome, identified using RF analysis. The features are ranked according to the mean decrease in accuracy the model suffers in sample classification, if the particular feature is removed from the analysis. The color codes indicate the association of each feature either with High PRADG or Low PRADG animals.

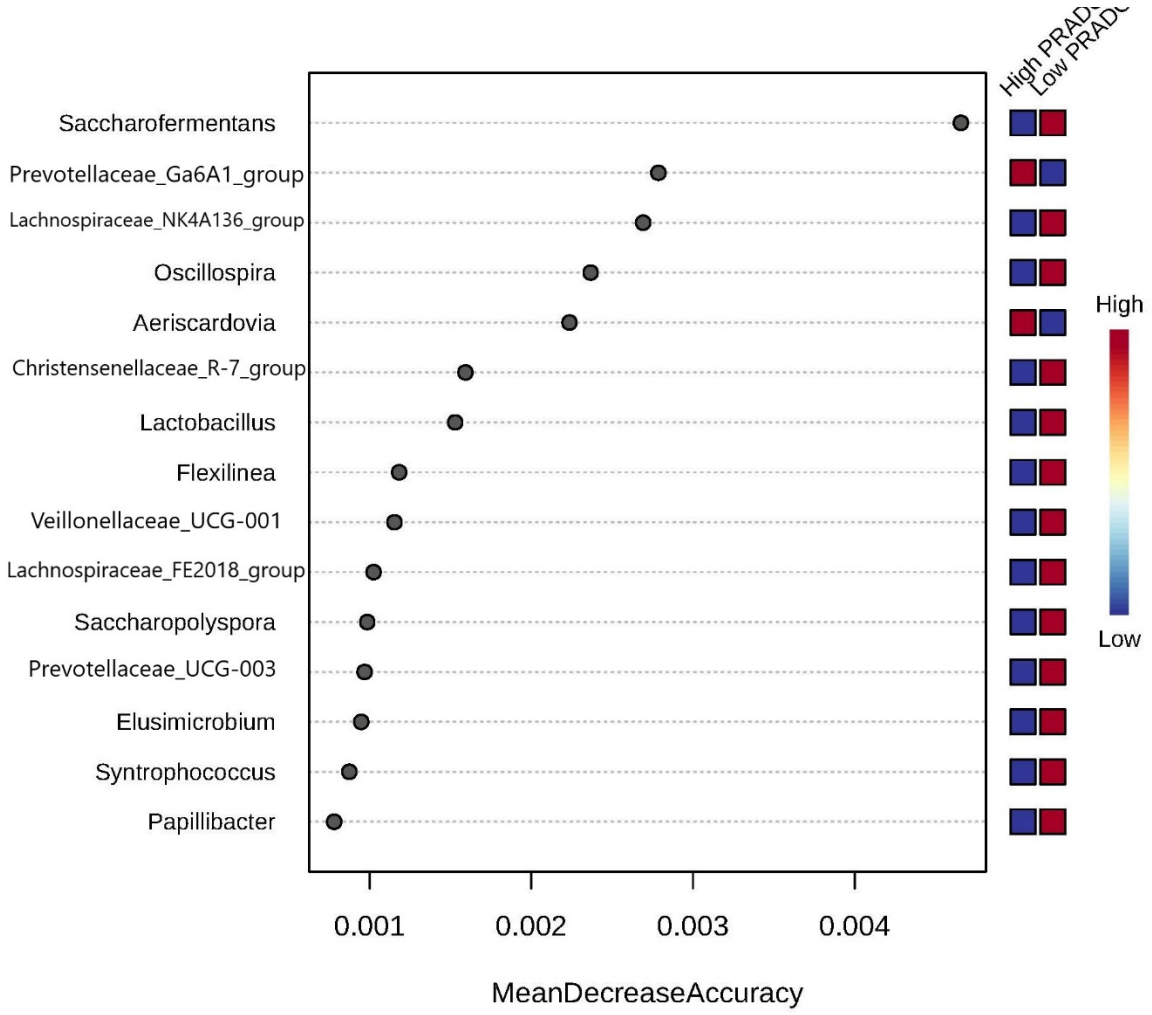


Figure 4.17: Illustration of differentially abundant taxa associated with High PRADG and Low PRADG animals identified by RF analysis of the restriction rumen samples. The features are ranked according to the mean decrease in accuracy the model suffers in assigning samples to each of the classes, if the particular feature is removed from the analysis. The color codes indicate the association of each feature either with High PRADG or Low PRADG animals.

4.4 Discussion

Often overlooked in animal nutrient requirement models, water is an essential ingredient in the diet of ruminants such as beef cattle, and plays a key role in all their essential physiological functions (Wickramasinghe et al. 2019). The relationship between water intake/impurities/temperature and performance in beef cattle has been well established in existing literature (Lofgreen et al., 1975; Lardner et al., 2005; Wright, 2007). The gut microbiome of cattle, commonly represented by its rumen and fecal microbiota is also well connected to animal performance (Myer, 2019). However, the association between the inherent gut microbiome of cattle, and animal performance in water restricted conditions, is yet to be extensively studied.

The biodiversity of any given microbial community is commonly assessed in terms of its alpha and beta diversity. The alpha diversity measures generated for both restriction rumen and restriction fecal samples obtained from high performing animals (High PRADG) did not yield any significant difference ($P>0.05$) when compared against the same from low performing (Low PRADG) animals. However, for alpha diversity indices that only take species richness into account, an increase in average alpha diversity was observed in High PRADG animals. This may be due to slight increase in the number of unique species present in fecal samples of animals that are performing better in water restricted conditions. However, for rumen samples, the alpha diversity of High PRADG animals decreased (even though not significant) when compared to Low PRADG animals (only in species richness indices). Further investigation at species level or strain level is

warranted to elucidate the cause behind this disparity in the behavior of rumen and fecal microbiomes during water restriction.

Similar behavior of the fecal and rumen microbiomes was observed during beta diversity analysis as well. Where the overall beta diversity of the fecal microbiome differed significantly between low performing and high performing animals, but no significant difference was observed in the rumen microbiome. The inherent differences in species diversity and composition that exist between rumen and fecal microbiomes (Seshadri et al., 2018; Lourenco et al., 2020) maybe the reason behind these differed responses of the rumen and fecal microbiome of cattle to water restriction`. Furthermore, the cumulative effect of reduced motility on cattle gut microbiome may be more visible in the fecal microbiome rather than the rumen microbiome, especially since the fecal microbiota may well be a combination of foregut and hindgut microbial communities of the animal.

In terms of microbial community composition and relative abundance of different taxa at genus level, *Christensenellaceae_R_7_group* demonstrated statistically significant ($P < 0.05$) increases in its relative abundances in the fecal and rumen microbiome of Low PRADG animals. Members of the *Christensenellaceae* family *(Christensenellaceae_R_7_group)* are known to increase in abundance when rumen pH decreases (De Nardi et al., 2016). The overall rumen pH of animals subjected to the current study showed a statistically significant decrease ($p=8.29e-05$), from a baseline average rumen pH of 6.45 to a restriction average rumen pH of 6.36. However, there was no significant difference ($p=0.073$) between the average rumen pH of low performing

(6.32) animals and high performing animals (6.51) during water restriction, but a slight decrease average pH was observed, thus strengthening the basis for our observation. Hence, this may be compelling evidence of an association between the high relative abundance of *Christensenellaceae_R_7_group* (driven by reduced overall rumen pH) and decreased performance of animals during water restriction.

Methanobrevibacter showed a statistically significant ($P<0.05$) increase in its abundance in the fecal microbiome of High PRADG animals under water restriction, when compared against Low PRADG animals. *Methanobrevibacter* also appears in the core microbiome High PRADG animals and gets replaced by *Familly_XIII_AD3011_group* in the Low PRADG animals. Increased abundance of methanogens such as *Methanobrevibacter* has been historically linked to high residual feed intake (HRFI), or low feed efficiency in cattle (Carberry et al., 2014; Delgado et al., 2019). However, no statistically significant association between RFI and performance parameters such as ADG has been found thus far in ruminants (Nkrumah et al., 2006; Zhang et al., 2017). Furthermore, a direct association between rumen/fecal methanogens, and ADG of cattle could not be found described in existing literature. Hence, further investigation is required to ascertain the association we observed between the abundance of *Methanobrevibacter* and the percentage recovery of ADG in feedlot cattle during water restriction.

The significant decrease ($P<0.05$) in the abundance of *Prevotella* in the rumen of Low PRADG animals supplements the behavior of *Christensenellaceae_R_7_group* and can

be considered as further evidence of increasing acidity in the rumen of low performing animals, since the decreasing abundance of *Prevotella* has been linked to pH drop in the rumen in cattle (Fernando et al., 2010; Holman and Gzyl, 2019).

The results from the biomarker analysis for fecal samples, also confirms the prominent presence of methanogens in High PRADG animals. Out of the 7 features identified by LEfSe analysis, the top 5 in terms of LDA scores contains two methanogens, namely; *Methanobrevibacter* and *Methanosphaera*. However, no methanogens appear as biomarkers that can differentiate High PRADG animals from Low PRADG animals under water restriction, in the rumen samples. Again, this agrees with our community composition and relative abundance data on restriction rumen samples, where *Methanobrevibacter* or any other methanogen does not show a statistically significant difference between High PRADG and Low PRADG animals.

Another noteworthy observation is, the genus *Akkermansia* being picked as a differentially abundant taxon in the fecal microbiome of Low PRADG animals during the biomarker analysis (by both LEfSe and RF). As described in the previous chapter, *Akkermansia* showed a statistically significant increase in abundance in the restriction fecal samples in comparison to baseline fecal samples. Building on that, in the current study we found that when compared to high performing animals (High PRADG) under water restriction, *Akkermansia* was differentially abundant in low performing animals (Low PRADG) and could be used as a biomarker to differentiate between them. Members of the genus *Akkermansia* are known to cause mucine degradation in the gut that causes

increased susceptibility to gut pathogens (Fan et al., 2020). Hence, this may be considered as evidence for gut health implications of water restriction in cattle that ultimately affects animal performance.

Members of the genus *Lactobacillus* have long been used as a probiotic in improving feed efficiency, growth and performance of animals in the cattle industry. However, while species such as *L. acidophilus, L. fermentum and L. ingluviei* are associated with weight gain in animals, other Lactobacilli such as *L. plantarum* and *L. gasseri* are known to cause weight loss (Million et al., 2012; Angelakis, 2017). Both Random Forest analysis and LEfSe analysis conducted during the current study identified *Lactobacillus* as a differentially abundant taxon in Low PRADG animals when compared to High PRADG animals. As mentioned above, these Lactobacilli present in Low PRADG animals can be species such as *L. plantarum* and *L. gasseri,* thus having an adverse effect on animal performance. However, it needs further investigation of fecal microbiome data at species level to ascertain this fact.

4.5 References

Ahlberg, C. M. 2018. Characterization of Water Intake in Beef Cattle: Test Length Guidelines, Water Intake Prediction, and Genetic Parameters. PhD Diss. Kansas State University, Manhattan.

Angelakis, E. 2017. Weight gain by gut microbiota manipulation in productive animals. Microb. Pathog. 106:162–170. doi:10.1016/j.micpath.2016.11.002.

Breiman, L. 2001. Random forests. Mach. Learn. 45:5–32. doi:10.1023/A:1010933404324.

Bruno, K. A. 2019. Effects of Extended Water Restriction On Performance, Behavior, Health, And Welfare of Feedlot Steers. PhD Diss. Oklahoma State University, Stillwater.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods. 13:581–583. doi:10.1038/nmeth.3869.

Carberry, C. A., S. M. Waters, D. A. Kenny, and C. J. Creevey. 2014. Rumen methanogenic genotypes differ in abundance according to host residual feed intake phenotype and diet type. Appl. Environ. Microbiol. 80:586–594. doi:10.1128/AEM.03131-13.

Chong, J., P. Liu, G. Zhou, and J. Xia. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat. Protoc. 15:799–821. doi:10.1038/s41596-019-0264-1.

Delgado, B., A. Bach, I. Guasch, C. González, G. Elcoso, J. E. Pryce, and O. Gonzalez-Recio. 2019. Whole rumen metagenome sequencing allows classifying and predicting feed efficiency and intake levels in cattle. Sci. Rep. 9:1–13. doi:10.1038/s41598-018-36673-w.

Dhariwal, A., J. Chong, S. Habib, I. L. King, L. B. Agellon, and J. Xia. 2017. MicrobiomeAnalyst: A web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. Nucleic Acids Res. 45:W180–W188. doi:10.1093/nar/gkx295.

Fan, P., B. Bian, L. Teng, C. D. Nelson, J. Driver, M. A. Elzo, and K. C. Jeong. 2020. Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation. ISME J. 14:302–317. doi:10.1038/s41396-019-0529-2.

Fernando, S. C., H. T. Purvis, F. Z. Najar, L. O. Sukharnikov, C. R. Krehbiel, T. G. Nagaraja, B. A. Roe, and U. De Silva. 2010. Rumen microbial population dynamics during adaptation to a high-grain diet. Appl. Environ. Microbiol. 76:7482–7490.

doi:10.1128/AEM.00388-10.

Holman, D. B., and K. E. Gzyl. 2019. A meta-analysis of the bovine gastrointestinal tract microbiota. FEMS Microbiol. Ecol. 95:1–9. doi:10.1093/femsec/fiz072..

Lardner, H. A., B. D. Kirychuk, L. Braul, W. D. Willms, and J. Yarotski. 2005. The effect of water quality on cattle performance on pasture. Aust. J. Agric. Res. 56:97–104. doi:10.1071/AR04086.

Lim, M. Y., E. J. Song, S. H. Kim, J. Lee, and Y. Do Nam. 2018. Comparison of DNA extraction methods for human gut microbial community profiling. Syst. Appl. Microbiol. 41:151–157. doi:10.1016/j.syapm.2017.11.008.

Lofgreen, G. P., R. L. Givens, S. R. Morrison, and T. E. Bond. 1975. Effect of Drinking Water Temperature on Beef Cattle Performance. J. Anim. Sci. 40:223–229. doi:10.2527/jas1975.402223x.

Lourenco, J. M., T. J. Kieran, D. S. Seidel, T. C. Glenn, M. F. Da Silveira, T. R. Callaway, and R. L. Stewart. 2020. Comparison of the ruminal and fecal microbiotas in beef calves supplemented or not with concentrate. PLoS One. 15:1–17. doi:10.1371/journal.pone.0231533.

Luz Calle, M. 2019. Statistical analysis of metagenomics data. Genomics and Informatics. 17. doi:10.5808/GI.2019.17.1.e6.

Million, M., E. Angelakis, M. Paul, F. Armougom, L. Leibovici, and D. Raoult. 2012. Comparative meta-analysis of the effect of Lactobacillus species on weight gain in humans and animals. Microb. Pathog. 53:100–108. doi:10.1016/j.micpath.2012.05.007.

Moore, S. G., A. C. Ericsson, S. E. Poock, P. Melendez, and M. C. Lucy. 2017. Hot topic: 16S rRNA gene sequencing reveals the microbiome of the virgin and pregnant bovine uterus. J. Dairy Sci. 100:4953–4960. doi:10.3168/jds.2017-12592.

Myer, P. R. 2019. Bovine Genome-Microbiome Interactions: Metagenomic Frontier for the Selection of Efficient Productivity in Cattle Systems. mSystems. 4:1–4. doi:10.1128/msystems.00103-19.

De Nardi, R., G. Marchesini, S. Li, E. Khafipour, K. J. C. Plaizier, M. Gianesella, R. Ricci, I. Andrighetto, and S. Segato. 2016. Metagenomic analysis of rumen microbial population in dairy heifers fed a high grain diet supplemented with dicarboxylic acids or polyphenols. BMC Vet. Res. 12:1–9. doi:10.1186/s12917-016-0653-4.

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

Ocejo, M., B. Oporto, and A. Hurtado. 2019. 16S rRNA amplicon sequencing characterization of caecal microbiome composition of broilers and free-range slow-growing chickens throughout their productive lifespan. Sci. Rep. 9:2506. doi:10.1038/s41598-019-39323-x.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41:590–596. doi:10.1093/nar/gks1219.

Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, and C. Huttenhower. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60. doi:10.1186/gb-2011-12-6-r60.

Seshadri, R., S. C. Leahy, G. T. Attwood, K. H. Teh, S. C. Lambie, A. L. Cookson, E. A. Eloe-Fadrosh, G. A. Pavlopoulos, M. Hadjithomas, N. J. Varghese, D. Paez-Espino, R. Perry, G. Henderson, C. J. Creevey, N. Terrapon, P. Lapebie, E. Drula, V. Lombard, E. Rubin, N. C. Kyrpides, B. Henrissat, T. Woyke, N. N. Ivanova, W. J. Kelly, N. Palevic, P. H. Janssen, R. S. Ronimus, S. Noel, P. Soni, K. Reilly, T. Atherly, C. Ziemer, A. D. Wright, S. Ishaq, M. Cotta, S. Thompson, K. Crosley, N. McKain, J. J. Wallace, H. J. Flint, J. C. Martin, R. J. Forster, R. J. Gruninger, T. McAllister, R. Gilbert, D. J. Ouwerkerk, A. J. Klieve, R. Al Jassim, S. Denman, C. McSweeney, C. Rosewarne, S. Koike, Y. Kobayashi, M. Mitsumori, T. Shinkai, S. Cravero, and M. Cerón Cucchi. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. Nat. Biotechnol. 36:359–367. doi:10.1038/nbt.4110.

Smith, B. 2011. Optimising Bacterial DNA Extraction from Faecal Samples: Comparison of Three Methods. Open Microbiol. J. 5:14–17. doi:10.2174/1874285801105010014.

Wickramasinghe, H. K. J. P., A. J. Kramer, and J. A. D. R. N. Appuhamy. 2019. Drinking water intake of newborn dairy calves and its effects on feed intake, growth performance, health status, and nutrient digestibility. J. Dairy Sci. 102:377–387. doi:10.3168/jds.2018-15579.

Williams, L. R., E. L. Jackson, G. J. Bishop-Hurley, and D. L. Swain. 2017. Drinking frequency effects on the performance of cattle: a systematic review. J. Anim. Physiol. Anim. Nutr. (Berl). 101:1076–1092. doi:10.1111/jpn.12640.

Wright, C. L. 2007. Management of Water Quality for Beef Cattle. Vet. Clin. North Am. - Food Anim. Pract. 23:91–103. doi:10.1016/j.cvfa.2006.12.002.

Zhang, X., W. Wang, F. Mo, Y. La, C. Li, and F. Li. 2017. Association of residual feed intake with growth and slaughtering performance, blood metabolism, and body composition in growing lambs. Sci. Rep. 7:1–11. doi:10.1038/s41598-017-13042-7.

CHAPTER V

POPULATION DYNAMICS OF THE BOVINE GUT MICROBIOME REVEALED BY METAGENOME SEQUENCING

5.1 Introduction

In the design of a microbiome study, 16S rRNA gene amplicon sequencing is often the first step in determining the structure and composition of the microbial community. Due to the continuous improvements in next generations sequencing (NGS) technologies, 16S rRNA gene amplicon studies have become cost effective and readily available. Hence, it has become the most widely used culture independent technique used to study microorganisms in their natural habitat (Gupta et al., 2019).

However, due to its inherent nature of only targeting a single or a few hypervariable regions within the 16S rRNA gene (the length of the entire gene at most), these studies are often incapable of accurate taxonomical profiling beyond genus level. Moreover, 16S rRNA gene amplicon sequence data lack the ability to provide any information about the functional or metabolic potential of the microbial community being studied. Other than that, like any other amplicon based study, they are prone to PCR bias, and any horizontal gene transfer events that have happened in the 16S locus between evolutionary distant

species may result in an overestimation of the true phylogenetic diversity (Sharpton, 2014; Rausch et al., 2019).

Therefore, whole genome shotgun sequencing (metagenome sequencing) is often utilized to go beyond traditional 16S rRNA gene amplicon sequencing and analyze the microbial community at species or strain level, while elucidating metabolic and functional potential of the microbiome. However, due to the comparatively high cost of whole genome shotgun sequencing, it is often difficult to use a large number of samples for this purpose. Hence, we utilized a tiered approach in the current study to select a subset of rumen and fecal samples used for the 16S gene amplicon study in order to have a deeper look at the structural and functional dynamics of the cattle gut microbiome under water restriction.

5.2 Materials and Methods

5.2.1 Animal Selection and DNA extraction

We selected 16 animals, out of the 146 animals used for the 16S rRNA gene amplicon study. Animal selection was done based on the beta diversity of their rumen and fecal samples, using the 'representative dissimilarity' method available in microPITA (Tickle et al., 2013). Attention was given to include a similar number of animals from each tail of the percentage recovery of average daily gain distribution. Sixty-four fecal and rumen samples originating from the 16 selected animals were subjected to whole genomic DNA extraction as described previously. These samples comprised of fecal and rumen samples from water intake and restriction trials for each animal. Molecular grade water was used

as the negative control, and ZymoBIOMICS® Microbial Community Standard (Zymo Research, Irvine, CA) was used as the positive control during DNA extraction.

5.2.2 Metagenomic sequencing

The extracted genomic DNA from all samples were sent to BGI Genomics (Cambridge, MA) for metagenome sequencing. Sequence library construction and shotgun metagenome sequencing was done using the BGISEQ-500 sequencing platform as described in Fang et al., 2018 to generate 40 million 150bp paired-end (PE) sequence reads worth 12Gb of data per sample.

5.2.3 Sequence data pre-processing and bioinformatics analysis

The raw data generated by the sequencing process was first filtered to remove adapter sequences and low quality sequences using SOAPnuke (Chen et al., 2018). Sequences that had a greater than 25% match with an adapter sequence was removed. Low quality sequence reads identified based on the presence of more than 50% of its bases with a quality score less than 20, were deleted. Sequence reads with more than 3% ambiguous bases, and duplicate sequences were also deleted.

The pre-processed, quality filtered sequences were then subjected to taxonomic classification using MetaPhlAn 3.0 (Truong et al., 2015). Both paired end sequencing data files in ‘fastq’ format obtained for each of the samples were utilized for the taxonomic analysis. MetaPhlAn 3.0 parameters such as ‘–t’, ‘--index’ and ‘--bt2_ps’ were set to ‘rel_ab_w_read_stats’ ‘latest’ and ‘sensitive’. ‘--add_viruses’ and ‘--

unknown_estimation' was also set to include viruses in the clade specific marker database search and to estimate the number of reads that do not map to known clades in the database. The taxonomic profiles generated for each sample were then combined into a single taxonomic profile similar to an ASV table where rows contain taxa and the columns would contain read counts for each sample. The combined taxonomic profile was then subjected to statistical analysis and data visualization using MicrobiomeAnalyst (Chong et al., 2020) and R.

HUMAnN 3.0 (Franzosa et al., 2018) was used to conduct functional profiling of the pre-processed and quality trimmed metagenome sequencing data. Both paired end sequencing files for each sample was subjected to analysis and default parameters were used to run HUMAnN 3.0. Functional profiles pertaining to abundance of gene families, abundance of metabolic pathways and pathway coverage were generated for metagenomic data obtained for each sample. Individual tables were joined to compile three comprehensive tables for gene families, metabolic pathways and pathway coverage for all fecal and rumen samples subjected to metagenome sequencing. Initial abundance values reported in Reads per Killobase (RPK) were normalized to Copies per Million (CPM) in order to facilitate comparison of samples with un-equal sequencing depth.

5.3 Results

5.3.1 Taxonomic profiling of rumen and fecal microbiomes at species level

5.3.1.1 Diversity analysis

We conducted alpha and beta diversity analyses (as described previously) at species level based on the taxonomic abundance profiles built using MetaPhlAn 3.0 for both baseline and restriction rumen samples and fecal samples. Chao1, observed species, Shannon and Simpson indices were used to measure alpha diversity, while Bray-Curtis dissimilarity and Jensen-Shannon index was used to measure beta diversity.

With regard to rumen samples, for the indices that take only species richness into account, a slight increase in average alpha diversity was observed in the restriction samples. For Shannon and Simpson indices that take both species richness and evenness into account, a slight decrease in the average alpha diversity was observed in restriction samples when compared to baseline samples (Figure 5.1). However, none of the between group comparisons (restriction against baseline) of alpha diversity at species level yielded statistically significant ($p<0.05$) results (Table 5.1).

The results of the beta diversity analysis for rumen samples followed a similar trajectory where, for both the beta diversity measures used, no clustering was observed when visualized using None-metric Multidimensional Scaling (NMDS) plots (Figure 5.2). A between group comparison of overall beta diversity using permutation analysis of variance (PERMANOVA) revealed no significant difference between the beta diversity of baseline and restriction rumen samples at species level (Table 5.2).

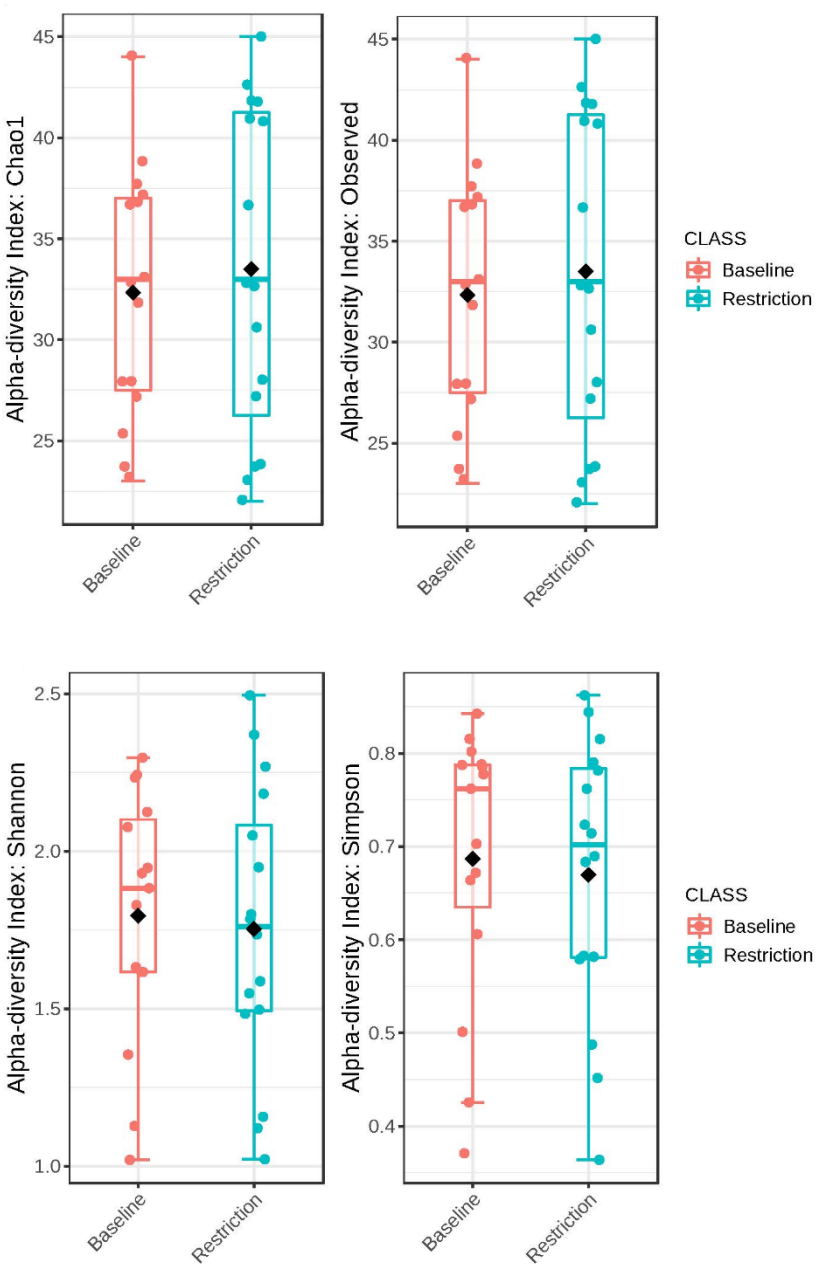


Figure 5.1: Alpha diversity box plots of baseline and restriction rumen samples based on the taxonomic profiles generated by MetaPhlAn 3.0. Y axis indicates the magnitude of the alpha diversity index calculated for each rumen sample. The samples are grouped and colored based on the water restriction or baseline trial they belong.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-0.9236	0.3635
Chao1	T-test	-0.4478	0.6577
Shannon	Man-Whitney	129	0.7404
Simpson	Man-Whitney	127	0.8000

Table 5.1: Between group comparison statistics for each of the alpha diversity indices used in estimating the alpha diversity of baseline and restriction rumen samples.

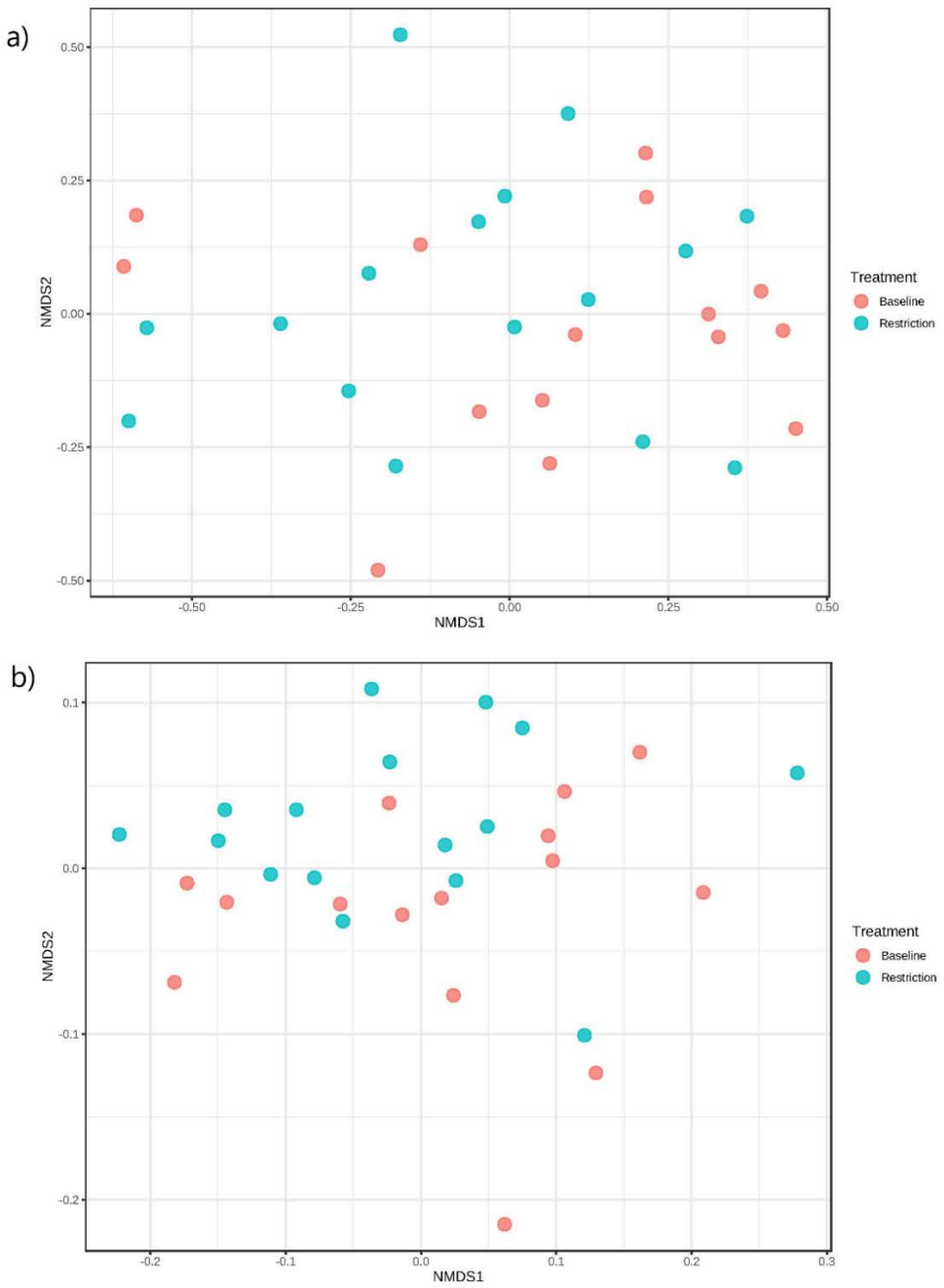


Figure 5.2: Non-metric multidimensional scaling (NMDS) plots illustrating the distribution of microbial communities from baseline and restriction rumen samples in terms of their beta diversity. Bray-Curtis dissimilarity index (a) and Jensen-Shannon divergence (b) were used as beta diversity estimates that measure dissimilarity between

samples based on the presence and absence of species. NMDS stress for plot (a) is 0.19099. NMDS stress for plot (b) is 0.1175.

Distance method	Statistical test	F-value	R-squared	p-value
Bray-Curtis dissimilarity	PERMANOVA	0.9422	0.0314	0.388
Jensen-Shannon divergence	PERMANOVA	0.9888	0.0329	0.379

Table 5.2: Between group comparisons of beta diversity measures using permutated analysis of variance (PERMANOVA).

We observed a similar pattern in the alpha diversity of fecal samples as well. Where a slight increase in species richness was seen in restriction samples for species richness indices (Chao1 and observed species), and a slight decrease was seen in restriction samples for indices that took both species richness and evenness into account (Figure 5.3). However as depicted in Table 5.3, none of these between group differences were statistically significant ($p<0.05$).

As depicted in Table 5.4, no significant difference could be observed between beta diversity of baseline and restriction fecal samples for either of the beta diversity measures used (Bray-Curtis dissimilarity and Jensen-Shannon divergence). Hence as expected, no clustering was observed in the NMDS plots as well (Figure 5.4).

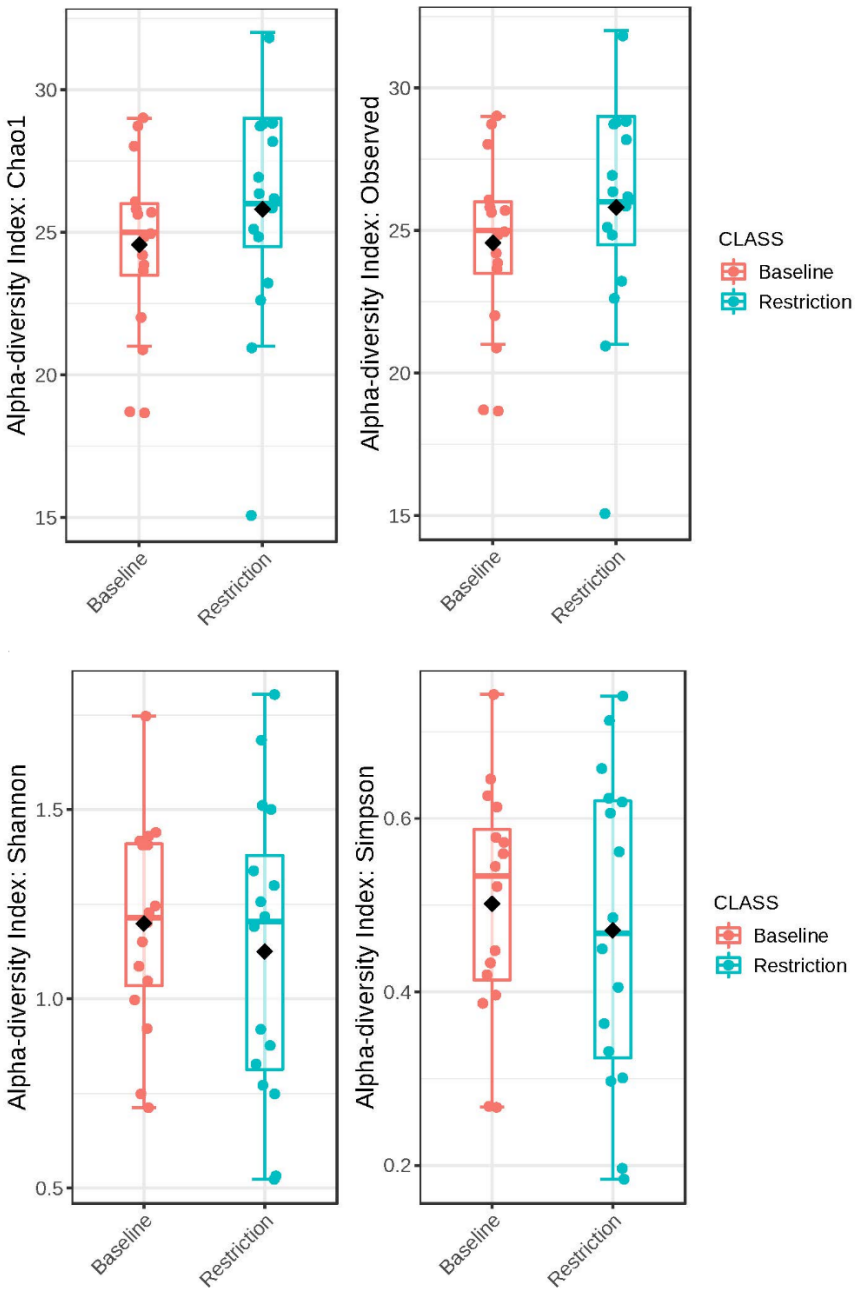


Figure 5.3: Box plots depicting species level alpha diversity distributions of baseline and restriction fecal samples. The alpha diversity values were derived based on taxonomic profiles generated using metagenome sequencing data. The value of the alpha diversity index calculated for each fecal sample is represented in the Y axis. The samples are grouped and colored based on the water restriction or baseline trial they originated from.

Alpha Diversity Index	Statistical Test	Test Statistic	p-value
Observed ASV	T-test	-0.9987	0.3216
Chao1	T-test	-0.9877	0.3317
Shannon	Man-Whitney	144	0.5641
Simpson	Man-Whitney	137	0.7520

Table 5.3: Results of the statistical tests conducted on each alpha diversity index to detect any significant differences in the alpha diversity that may exist between baseline and restriction fecal samples.

Distance method	Statistical test	F-value	R-squared	p-value
Bray-Curtis dissimilarity	PERMANOVA	1.3838	0.0440	0.210
Jensen-Shannon divergence	PERMANOVA	2.6839	0.0821	0.054

Table 5.4: Results of the statistical test (PERMANOVA) conducted to compare the overall beta diversity of baseline and restriction fecal samples

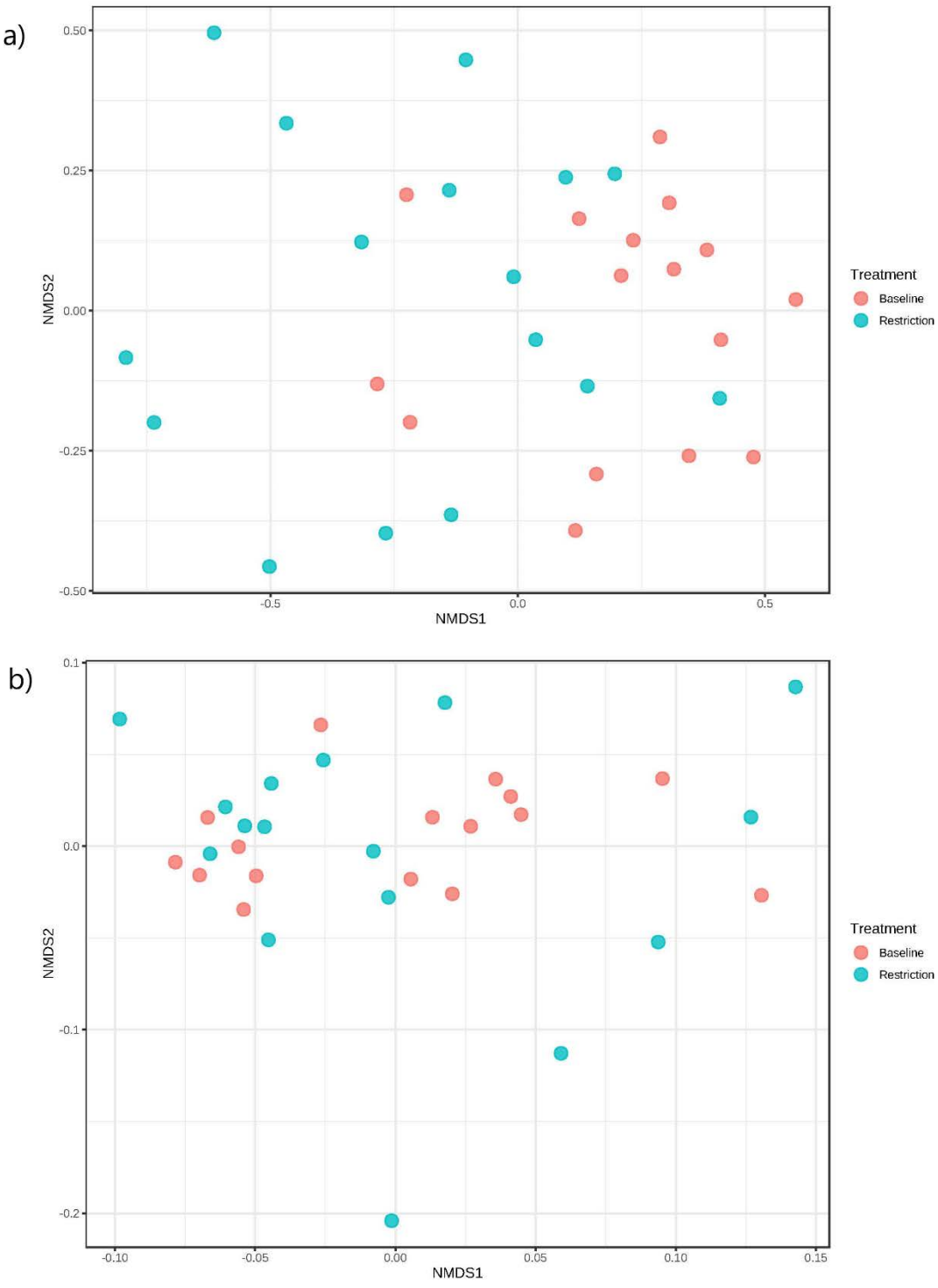


Figure 5.4: Non-metric multidimensional scaling (NMDS) plots illustrating the distribution of baseline and restriction fecal samples based on Bray-Curtis (a) and Jensen-

Shannon (b) dissimilarities. NMDS stress for plot (a) was 0.2062 and for plot (b) it was 0.1582.

5.3.1.2 Abundance Profiles

Using MetaPhlAn 3.0, we were able to accurately map an estimated 86,644,338 sequence reads (out of an estimated total of 2,726,200,094 sequence reads) from rumen samples to its clade specific marker gene database. The mapped sequence reads contained a total 130 species belonging to bacteria, archaea and viruses. Similarly, a total of 131,259,912 (out of an estimated total of 2,939,333,876) sequence reads arising from the fecal samples were mapped to 152 species in the marker gene database. The subsequent taxonomic profiles for all rumen and fecal samples were built based on taxonomic classification of the mapped reads.

As depicted in Figure 5.5 and Figure 5.6, the 10 most predominant organisms identified at species level for both baseline and restriction rumen samples were *Bacteroidales bacterium* KHT7, *Methanobrevibacter millerae, Methanosphaera sp* BMS, *Prevotella ruminicola, Prevotella sp* tf2 5, *Pseudobutyrivibrio ruminis, Ruminococcaceae bacterium* P7, *Sarcina sp* DSM 11001, *Streptococcus equinus,* and *Succiniclasticum ruminis.* A notable increase in abundance was observed in *Sarcina sp* DSM 11001 from baseline (41%) to restriction (47%) rumen samples. For *Succiniclasticum ruminis, Prevotella ruminicola* and *Prevotella sp* tf2 5, a 5.0%, 2.0% and 3.5% decrease in relative abundance was observed in rumen restriction samples compared to baseline samples. However, these fluctuations in relative abundance were not found to be statistically significant ($p<0.05$).

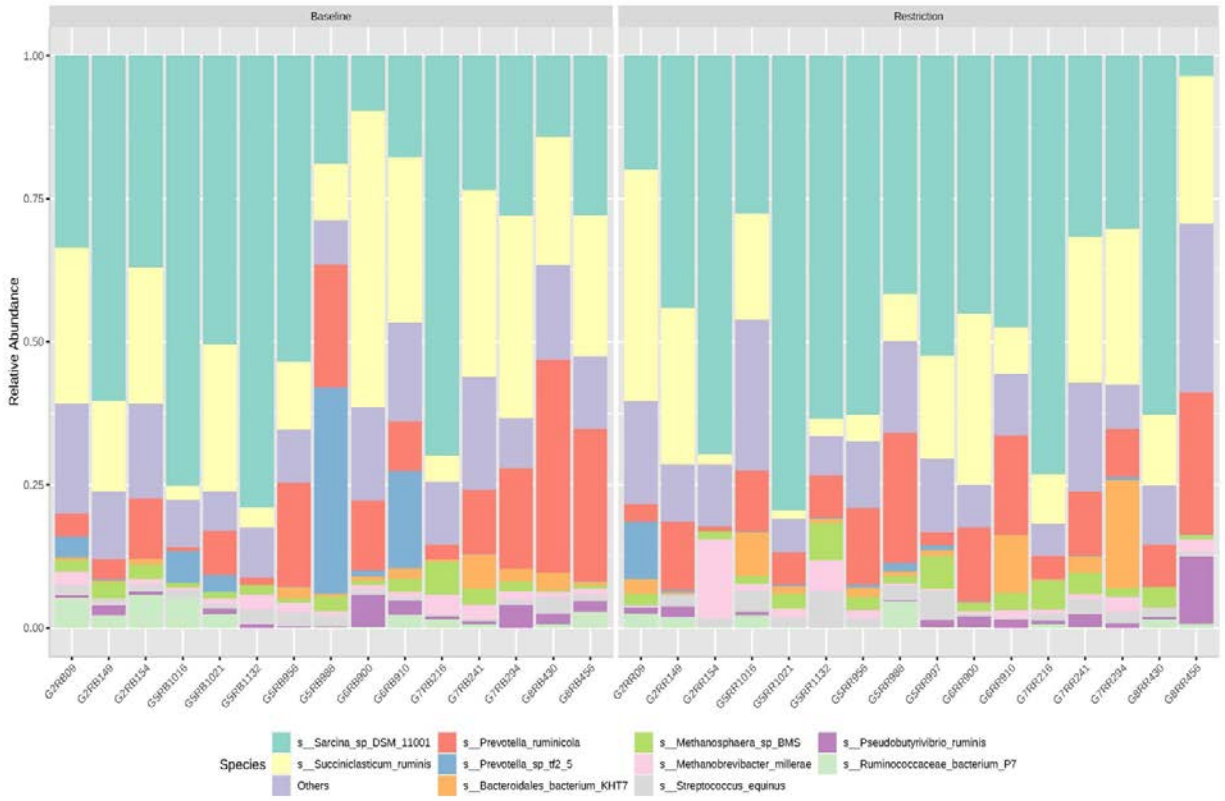


Figure 5.5: Animal wise relative abundance of microbial taxa in baseline and restriction rumen samples at species level. The 10 most predominant species identified in both baseline and restriction samples are depicted. The bar plot was created by summarizing taxonomic profiling data generated using MetaPhlAn 3.0 at species level.

Species	**Baseline**	**Restriction**	**P-value**
Sarcina sp DSM 11001	40.0%	47.2%	0.27
Succiniclasticum ruminis	21.3%	16.2%	0.08
Prevotella ruminicola	12.2%	10.2%	0.60
Prevotella sp tf2 5	4.4%	0.9%	0.13
Bacteroidales bacterium KHT7	1.4%	3.1%	0.22
Methanosphaera sp BMS	1.8%	2.6%	0.29
Methanobrevibacter millerae	1.4%	2.1%	0.33
Streptococcus equinus	1.3%	2.0%	0.09
Pseudobutyrivibrio ruminis	1.5%	1.6%	0.94
Ruminococcaceae bacterium P7	1.9%	0.9%	0.20

Table 5.5: The change in the relative abundance of the 10 most predominant species in baseline and restriction rumen samples and the p-values calculated based on animal wise fluctuations of each species.

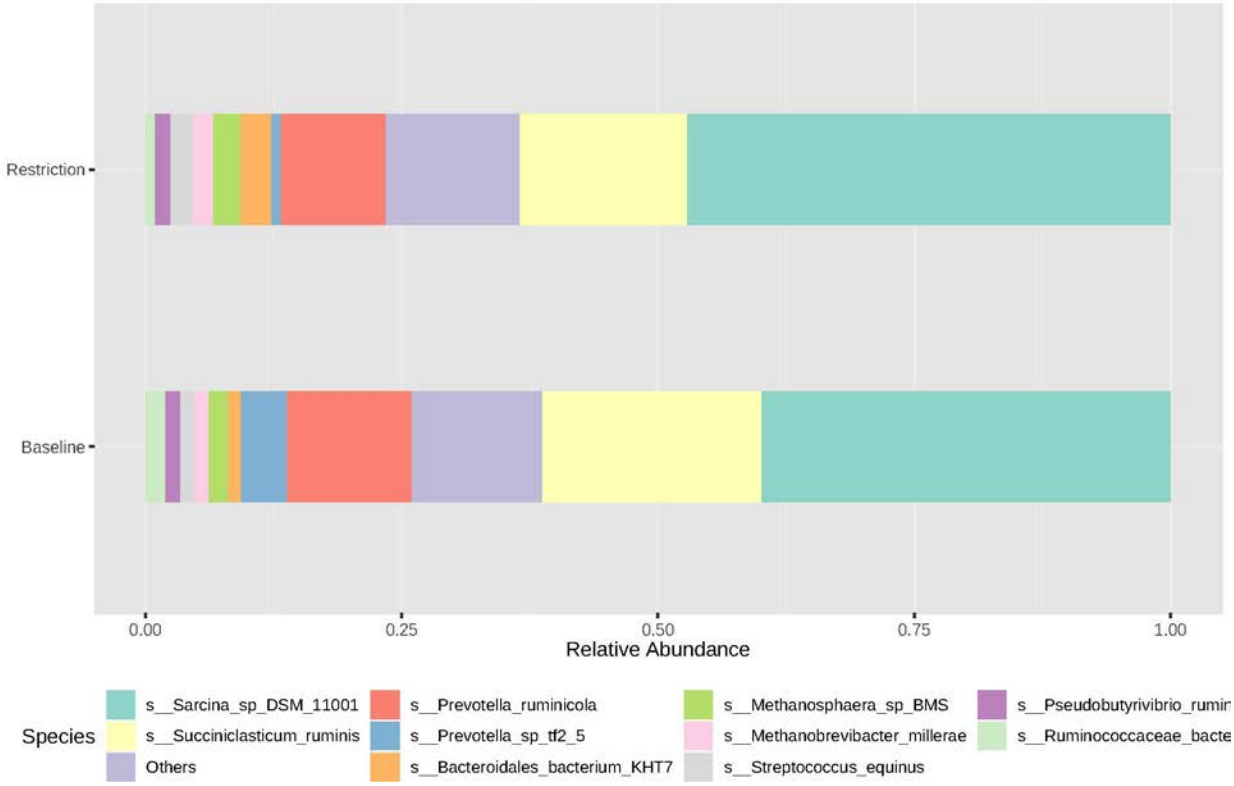


Figure 5.6: The relative abundance of the 10 most predominant species in baseline and restriction rumen samples summarized by treatment.

As depicted in Figure 5.7 and 5.8 the 10 most predominant species in baseline and restriction fecal samples essentially remained the same, and comprised of *Sarcina sp* DSM 11001, *Turicibacter sanguinis*, *Ruminococcaceae bacterium* P7, *Methanosphaera*

sp BMS, *Streptococcus equinus*, *Escherichia coli*, *Methanobrevibacter thaueri*, *Methanobrevibacter millerae*, *Saccharopolyspora rectivirgula*, and *Treponema porcinum.* However, as depicted in Table 5.6, a significant decrease in the abundance of *Ruminococcaceae bacterium P7*, *Methanosphaera sp BMS* and *Methanobrevibacter millerae* was observed in restriction fecal samples, when compared to the baseline.

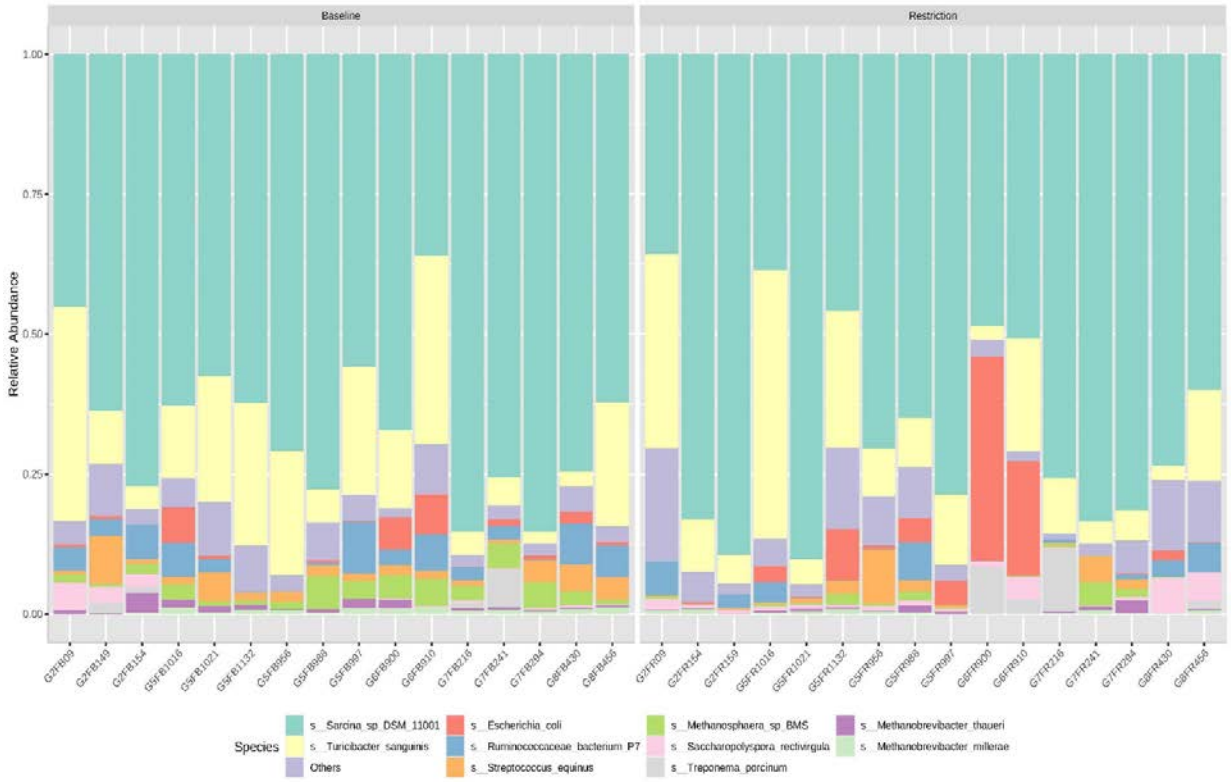


Figure 5.7: Animal to animal variation in relative abundance of the 10 most predominant species identified in baseline and restriction fecal samples. The relative abundance values are derived from taxonomic profiles built using metgenome sequencing data obtained from baseline and restriction fecal samples selected for the study.

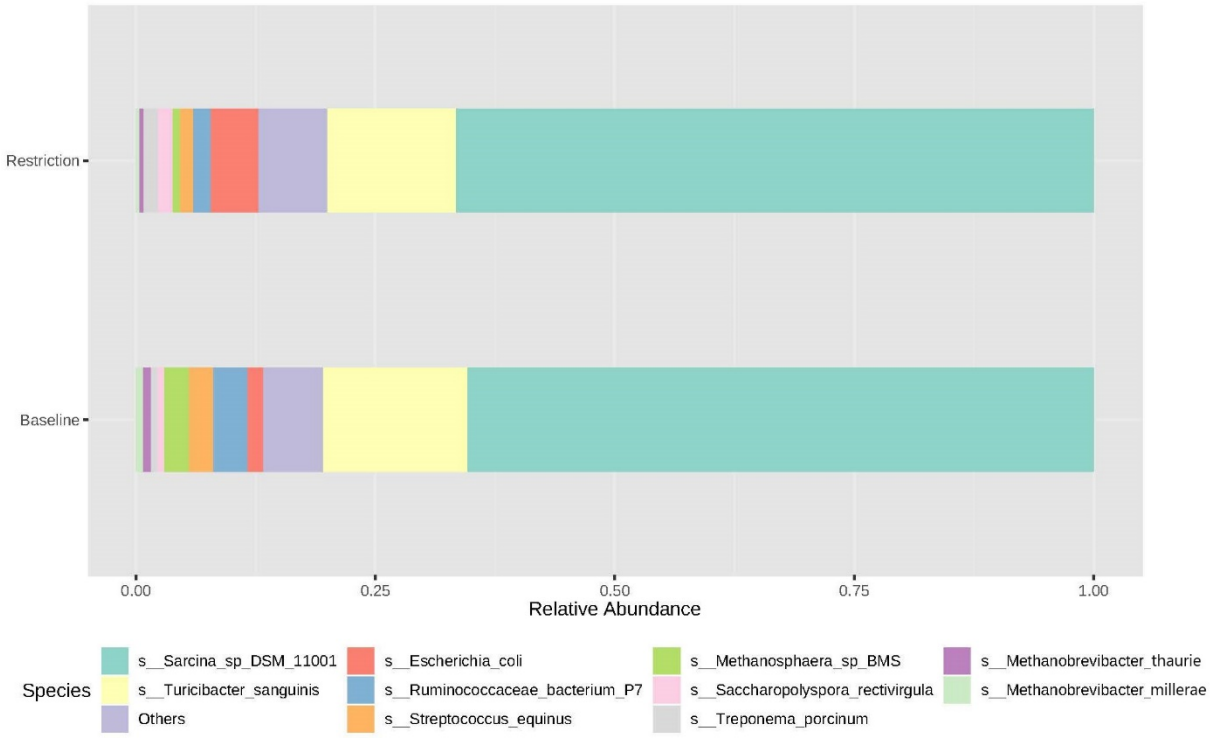


Figure 5.8: The relative abundance of the 10 most predominant species found in baseline and restriction fecal samples summarized by treatment.

Species	**Baseline**	**Restriction**	**P-value**
Sarcina sp DSM 11001	65.39%	66.62%	0.7597
Turicibacter sanguinis	15.04%	13.37%	0.5931
Ruminococcaceae bacterium P7*	3.64%	1.81%	0.0375
*Methanosphaera sp BMS***	2.68%	0.82%	0.0006
Streptococcus equinus	2.42%	1.35%	0.2534
Escherichia coli	1.61%	5.01%	0.1207
Methanobrevibacter thaueri	0.84%	0.42%	0.1765
*Methanobrevibacter millerae**	0.73%	0.41%	0.0186
Saccharopolyspora rectivirgula	0.68%	1.49%	0.0897
Treponema porcinum	0.64%	1.47%	0.3991

(*) – $p<0.05$; (**) – $p<0.001$

Table 5. 6: The 10 most abundant species observed in the baseline and restriction fecal samples and their respective relative abundances. P-values indicated are calculated based on animal wise fluctuations of the microorganisms between baseline and restriction trials.

5.3.1.3 Core microbiome analysis

The taxonomic profiles generated using MetaPhlAn 3.0 (Truong et al., 2015) were analyzed using MicrobiomeAnalyst (Chong et al., 2020) to derive rumen and fecal core microbiome profiles. As described previously, a relative abundance > 0.01% and prevalence > 70% were used as selection criteria when defining the core microbiomes. Figure 5.5, depicts the core microbiomes derived from baseline and restriction rumen samples at species level.

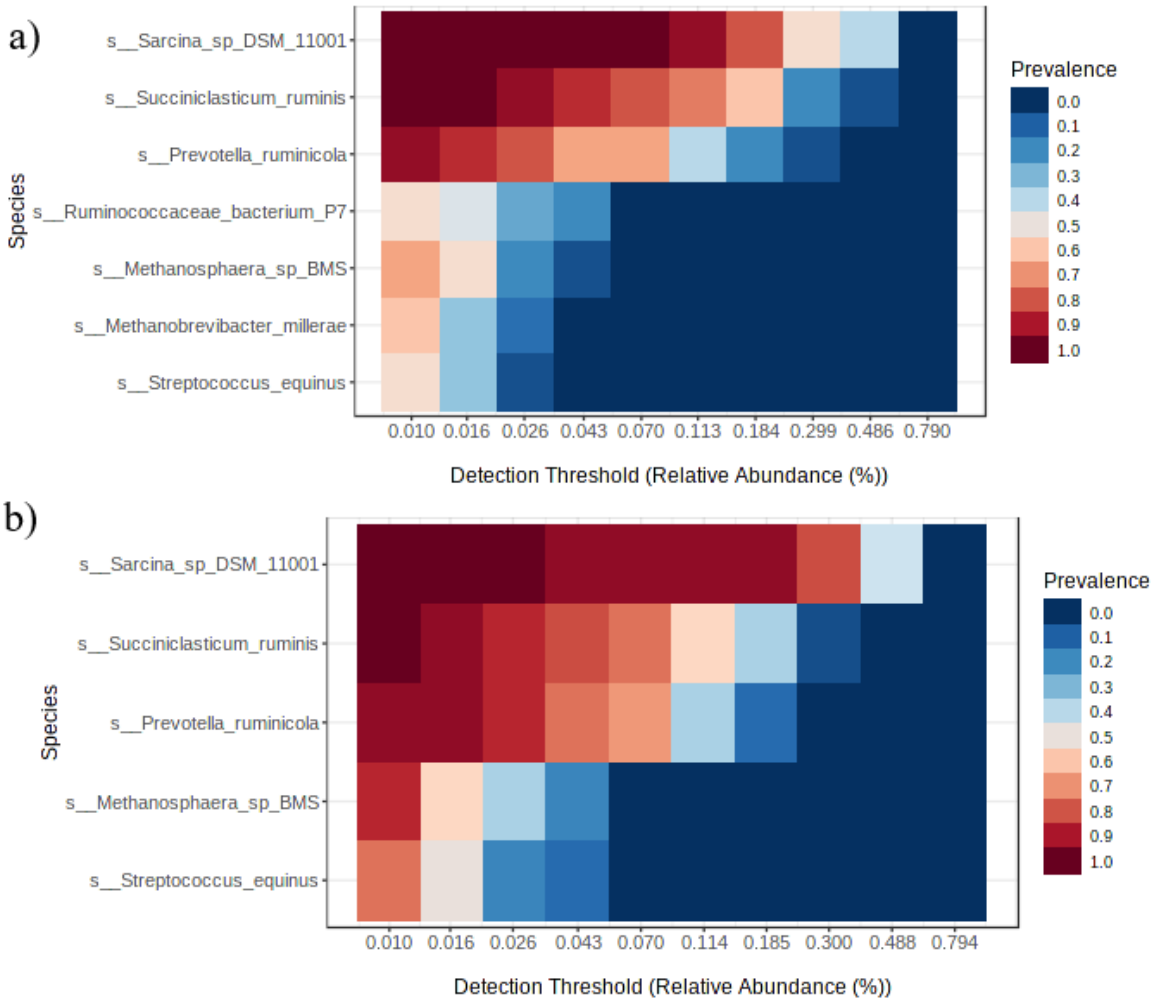


Figure 5.9: The core microbiomes of baseline (a) and restriction (b) rumen samples at species level, based on the taxonomic profiling of shotgun metagenome sequencing data. A relative abundance greater than 0.01% and prevalence greater than 70% was used as the cut-off values for a taxon to be included in the particular core microbiome.

The baseline core microbiome for rumen samples comprised of 7 species, namely; *Sarcina sp DSM 11001, Succniclasticum ruminis, Prevotella ruminicola, Ruminococcaceae bacterium P7, Methanosphaera sp BMS, Methanobrevibacter*

millerae, and Streptococcus equinus. The core microbiome of restriction rumen samples constituted of only five species, namely; *Sarcina sp* DSM 11001, *Succiniclasticum ruminis, Prevotella ruminicola, Methanosphaera sp* BMS, *and Streptococcus equinnus.* Hence, the rumen core microbiome identified here at species level appears to shrink during water restriction, losing two of its members, *Ruminococcaceae bacterium* P7 and *Methanobrevibacter millerae*.

Sarcina sp. DSM 11001, *Turicibacter sanguinis, Ruminococcaceae bacterium* P7, *Methanosphera sp* BMS and *Streptococcus equinus* formed the baseline fecal microbiome. However, during water restriction the species level fecal core microbiome of the animals used for the study appears to shrink more than 50% in size (in terms of number of species), comprising only of *Sarcina sp*. DSM 11001 and *Turicibacter sanguinis.*

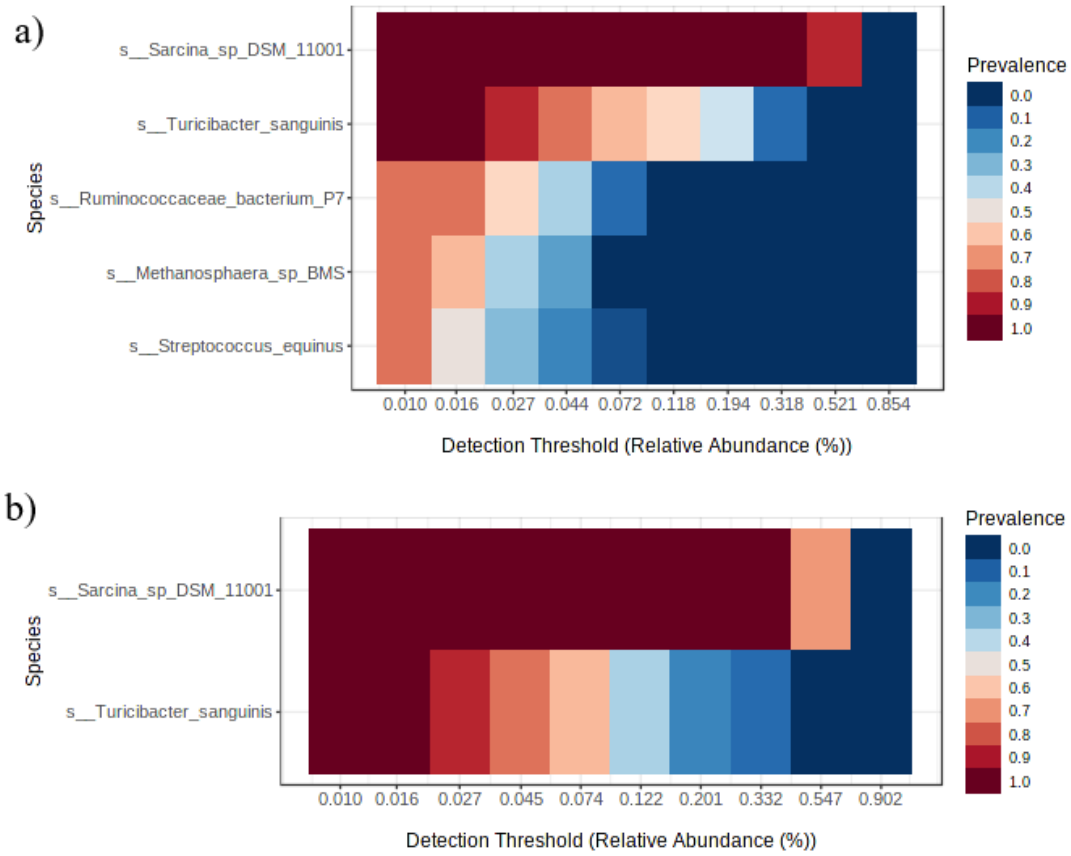


Figure 5.10: Species level core microbiomes of baseline (a) and restriction (b) fecal samples derived via taxonomic profiling of shotgun metagenome sequencing data. A relative abundance greater than 0.01% and prevalence greater than 70% was used as cut-off values for defining the composition of the core microbiomes.

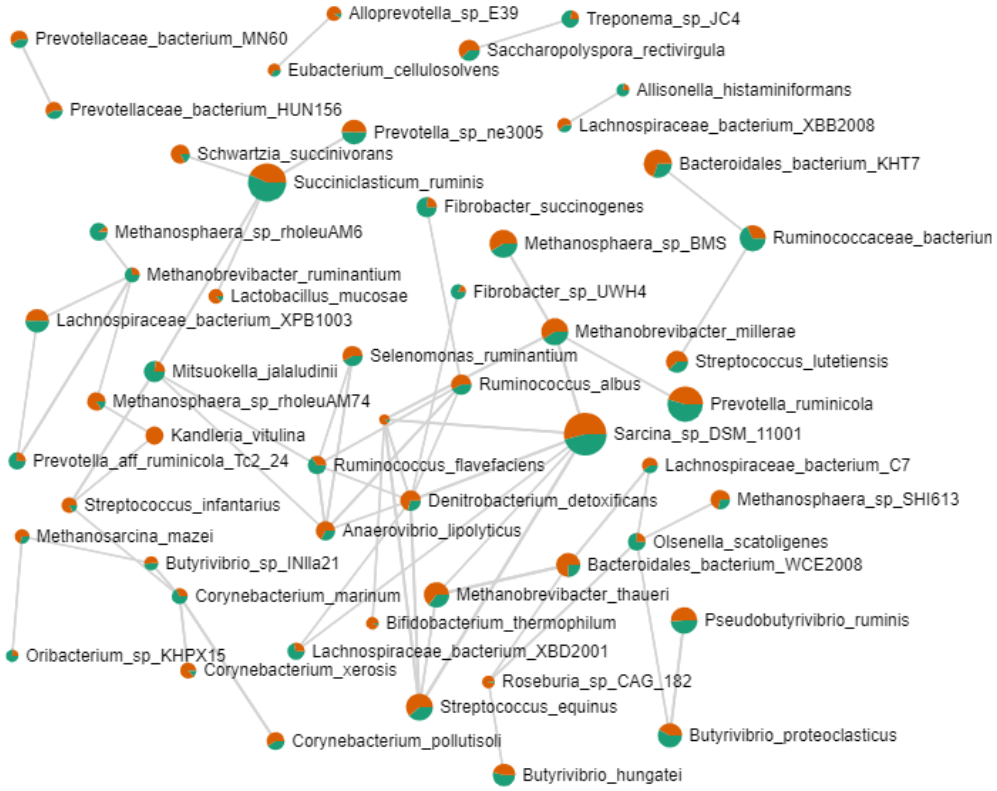
5.3.1.4 Co-occurrence networks

We constructed microbial co-occurrence networks using baseline and restriction rumen and fecal samples in order to visualize possible interactions between the cattle gut microorganisms, and to identify organisms that may act as potential keystone species in their response to water restriction. MicrobiomeAnalyst (Chong et al., 2020) was used to construct and visualize the networks, based on the taxonomic profiles built by

MetaPhlAn 3.0. The SparCC algorithm (Friedman and Alm, 2012; Chong et al., 2020) was used to calculate pairwise correlations between species. A p-value of 0.05 and a correlation of 0.6 was used as thresholds when generating the networks. As depicted in Figure 5.11, two co-occurrence networks were built, each for rumen and fecal samples, based on the co-variation of species abundance during water restriction.

Upon visual examination of the co-occurrence network generated for rumen samples, species such as *Sarcina sp DSM 1101, Denitrobacterium detoxificans, Ruminococcus flavefaciens, Streptococcus equinus* and *Methanobravibactor thaueri* appear to act as hubs in the network with many significant correlations to other species and subnetworks. In the fecal samples, species such as *Geobacillus thermodenitrificans, Streptococcus lutetiencis, Streptococcus equinus,* and *Methanobrevibactor millerae* appears to act as hubs of the co-occurrence network.

a)


Alloprevotella_sp_E39
Treponema_sp_JC4
Prevotellaceae_bacterium_MN60
Saccharopolyspora_rectivirgula
Eubacterium_cellulosolvens
Allisonella_histaminiformans
Prevotellaceae_bacterium_HUN156
Lachnospiraceae_bacterium_XBB2008
Prevotella_sp_ne3005
Schwartzia_succinivorans
Bacteroidales_bacterium_KHT7
Succiniclasticum_ruminis
Fibrobacter_succinogenes
Methanosphaera_sp_rholeuAM6
Methanosphaera_sp_BMS
Ruminococcaceae_bacterium
Methanobrevibacter_ruminantium
Lactobacillus_mucosae
Fibrobacter_sp_UWH4
Lachnospiraceae_bacterium_XPB1003
Methanobrevibacter_millerae
Selenomonas_ruminantium
Streptococcus_lutetiensis
Mitsuokella_jalaludinii
Ruminococcus_albus
Methanosphaera_sp_rholeuAM74
Prevotella_ruminicola
Sarcina_sp_DSM_11001
Kandleria_vitulina
Prevotella_aff_ruminicola_Tc2_24
Ruminococcus_flavefaciens
Lachnospiraceae_bacterium_C7
Streptococcus_infantarius
Denitrobacterium_detoxificans
Methanosphaera_sp_SHI613
Methanosarcina_mazei
Anaerovibrio_lipolyticus
Olsenella_scatoligenes
Butyrivibrio_sp_INIla21
Bacteroidales_bacterium_WCE2008
Corynebacterium_marinum
Methanobrevibacter_thaueri
Bifidobacterium_thermophilum
Pseudobutyrivibrio_ruminis
Oribacterium_sp_KHPX15
Lachnospiraceae_bacterium_XBD2001
Corynebacterium_xerosis
Roseburia_sp_CAG_182
Streptococcus_equinus
Corynebacterium_pollutisoli
Butyrivibrio_proteoclasticus
Butyrivibrio_hungatei

b)

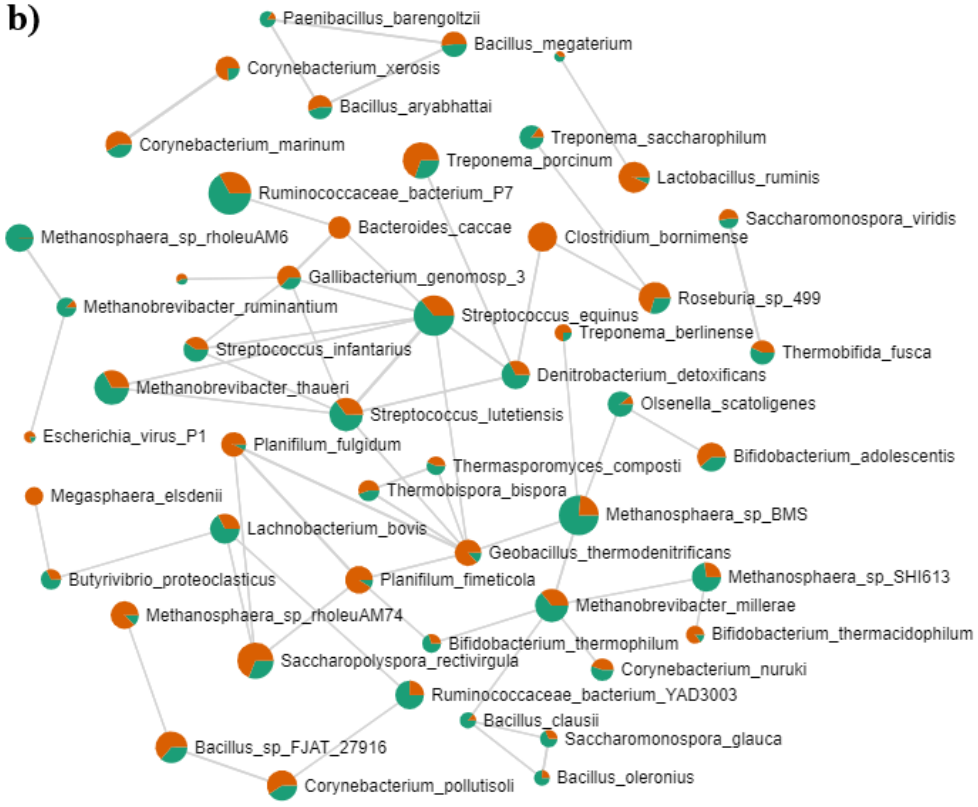

Paenibacillus_barengoltzii
Bacillus_megaterium
Corynebacterium_xerosis
Bacillus_aryabhattai
Corynebacterium_marinum
Treponema_saccharophilum
Treponema_porcinum
Lactobacillus_ruminis
Ruminococcaceae_bacterium_P7
Saccharomonospora_viridis
Methanosphaera_sp_rholeuAM6
Bacteroides_caccae
Clostridium_bornimense
Gallibacterium_genomosp_3
Roseburia_sp_499
Methanobrevibacter_ruminantium
Streptococcus_equinus
Treponema_berlinense
Streptococcus_infantarius
Thermobifida_fusca
Denitrobacterium_detoxificans
Methanobrevibacter_thaueri
Olsenella_scatoligenes
Streptococcus_lutetiensis
Escherichia_virus_P1
Planifilum_fulgidum
Bifidobacterium_adolescentis
Thermasporomyces_composti
Thermobispora_bispora
Megasphaera_elsdenii
Methanosphaera_sp_BMS
Lachnobacterium_bovis
Geobacillus_thermodenitrificans
Butyrivibrio_proteoclasticus
Planifilum_fimeticola
Methanosphaera_sp_SHI613
Methanobrevibacter_millerae
Methanosphaera_sp_rholeuAM74
Bifidobacterium_thermophilum
Bifidobacterium_thermacidophilum
Saccharopolyspora_rectivirgula
Corynebacterium_nuruki
Ruminococcaceae_bacterium_YAD3003
Bacillus_clausii
Saccharomonospora_glauca
Bacillus_sp_FJAT_27916
Bacillus_oleronius
Corynebacterium_pollutisoli

Figure 5.11: Microbial co-occurrence networks derived for rumen (a) and fecal (b) microbiomes considered in the study. SparCC algorithm was used to calculate pairwise correlations. A p-value < 0.05 and a correlation threshold > 0.6 was used to identify an interaction between two organisms. The edges represent negative or positive correlations between the two connecting organisms (nodes). The node sizes correspond to the abundance of each organism in restriction and baseline samples. Orange represents the mean abundance of the given organism in restriction samples and green indicates the same in baseline samples.

5.3.1.5 Differentially abundant microorganisms

In order to identify differentially abundant microorganisms at specie level, we conducted a LEfSe (linear discriminant analysis effect size) (Segata et al., 2011) analysis on the taxonomic profiles built using metagenome sequencing data. As described previously, we used a

p-value <0.05 and an LDA (linear discriminant analysis) score > 2.0 when selecting features. Two separate biomarker profiles were generated for rumen and fecal samples as depicted in Figure 5.12.

Through the LEfSe analysis we were able to identify 8 differentially abundant species for rumen samples and 12 differentially abundant species for fecal samples. Species such as *Kandleria vitulina, Denitrobacterium detoxificans, Schwartzia succinivorans, Saccharopolyspora rectivirgula,* and *Streptococcus infantarius* with known gut health implications were found to be enriched in restriction rumen samples.

Notably, in the fecal samples *Methanosphera sp. BMS, Methanobrevibacter thaueri,* and *Methanobrevibacter millerae* that together forms a major part of the population of methanogens in cattle, were all enriched in the baseline samples in comparison to restriction samples. Similar to their behavior observed in the rumen samples, *Saccharopolyspora rectivirgula* could be seen enriched in the restriction fecal samples as well. *Geobacillus thermodentrificans* and *Roseburia sp.* 499 are the other notable species observed to be enriched in the fecal samples collected during water restriction.

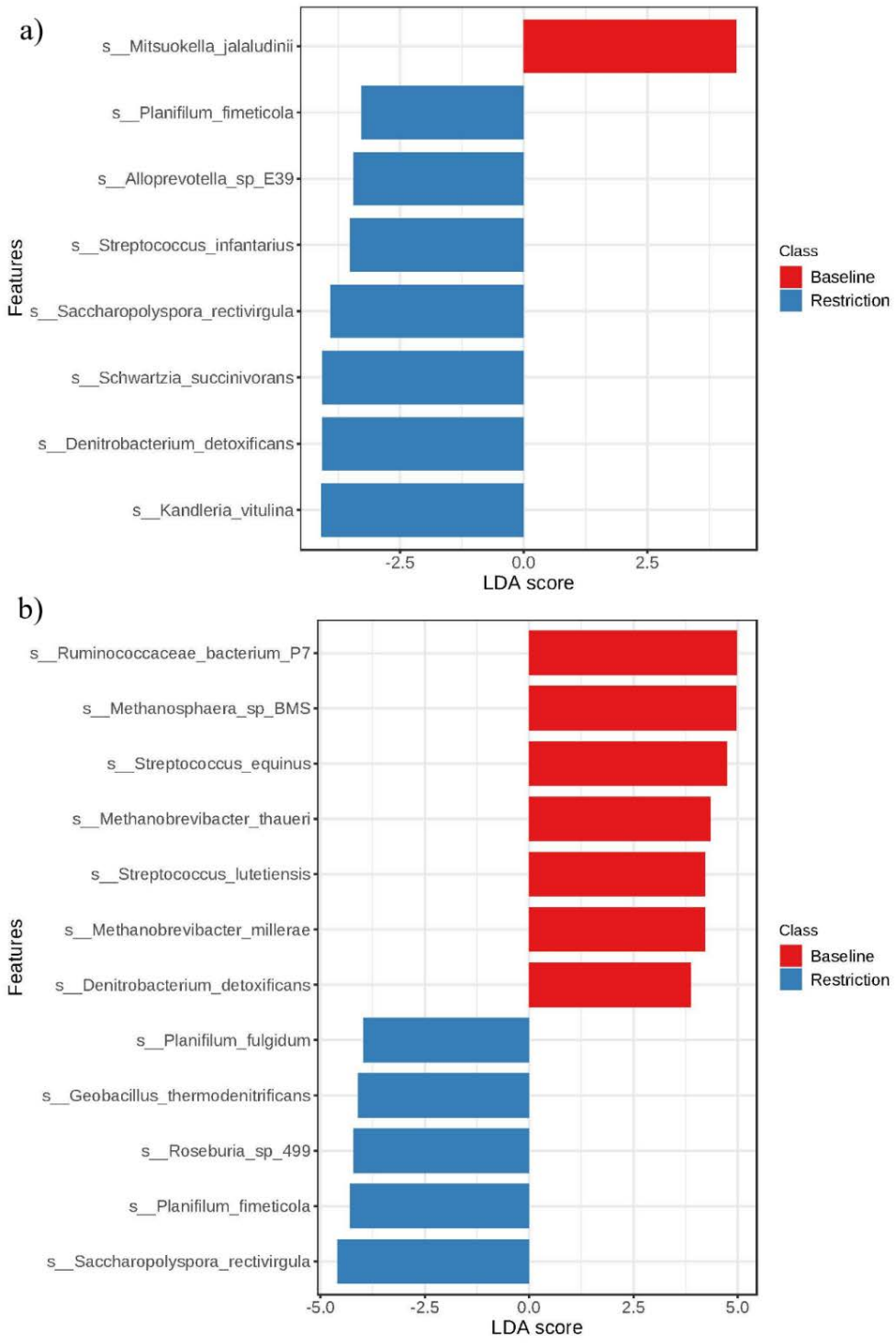


Figure 5. 12: Differentially abundant species identified in the rumen (a) and fecal (b) samples based on their enrichment in the respective restriction and bassline samples. A p-value < 0.05 and an LDA score > 2.0 was used to classify species into their enriched

groups (baseline and restriction). The red bars represent the LDA score of microbial species enriched in baseline samples and the blue bars indicate the LDA score of species that are enriched in restriction samples.

5.3.2 Functional dynamics of the beef cattle gut microbiome during water restriction

In order to unravel the changes that occur in the functional profile of the rumen and fecal microbiomes during water restriction, we used HUMAnN 3.0 (Franzosa et al., 2018) to detect the presence/absence and the abundance of metabolic pathways in the rumen and fecal microbiomes. Quality controlled metagenome sequencing data generated from selected rumen and fecal samples (as described previously) obtained during the baseline and water restriction trials were used for this purpose. STAMP 2.1.3 (Parks et al., 2014) was used for statistical analysis and visualization of pathway abundance data generated using HUMAnN 3.0.

5.3.2.1 Functional dynamics of the rumen microbiome during water restriction

We were able to identify 90 microbial metabolic pathways that had a statistically significant ($p<0.05$) difference in their abundance when compared between baseline and restriction rumen samples. Biologically important metabolic pathways such as (but not limited to) glucose-1-phosphate degradation, glycogen biosynthesis I (from ADP-D-glucose), homolactic fermentation, TCA cycle, super pathway of purine nucleotide salvage, pyrimidine nucleobases salvage, thiamin salvage, thiamin diphosphate biosynthesis, pyrimidine deoxyribonucleotide de novo biosynthesis, L-histidine biosynthesis, and super pathway of L-lysine, L-threonine and L-methionine biosynthesis

had significantly high ($p<0.05$) pathway abundance in restriction rumen samples (Figure 5.13).

Furthermore, metabolic pathways such as superpathway of branched chain amino acid biosynthesis, L-methionine biosynthesis III, L-isoleucine biosynthesis (I and II), L-lysine biosynthesis III, pyruvate fermentation to isobutanol, adenosine ribonucleotides de novo biosynthesis, purine ribonucleosides degradation, and superpathway of L-serine and glycine biosynthesis I were observed to be significantly low in abundance in the restriction rumen samples (Figure 5.14).

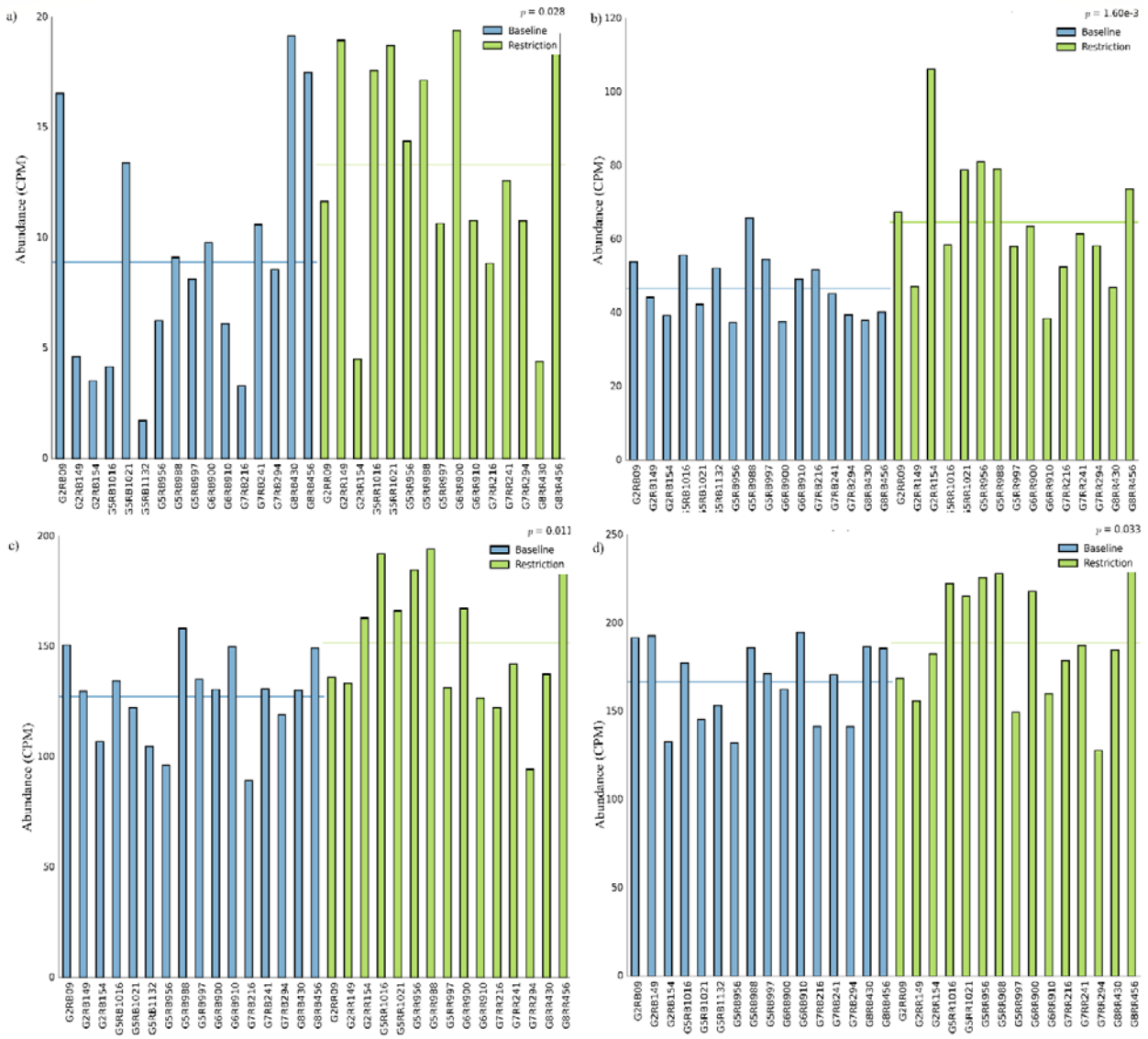


Figure 5.13: A subset of the pathways that were found to be significantly ($p<0.05$) high in abundance in restriction rumen samples; a) Glucose-1-phosphate degradation ($p=0.028$), b) Glycogen biosynthesis from ADP-D-Glucose ($p=1.6e-3$), c) Homolactic fermentation ($p=0.01$), d) super pathway of L-lysine, L-threonine and L-methionine biosynthesis ($p=0.03$). Baseline rumen samples are indicated in blue and restriction rumen samples are indicated in green. Mean pathway abundance for each group is

indicated by the horizontal line. The Y-axis indicates pathway abundance in counts per million (CPM) for each sample and the X-axis indicates samples IDs.

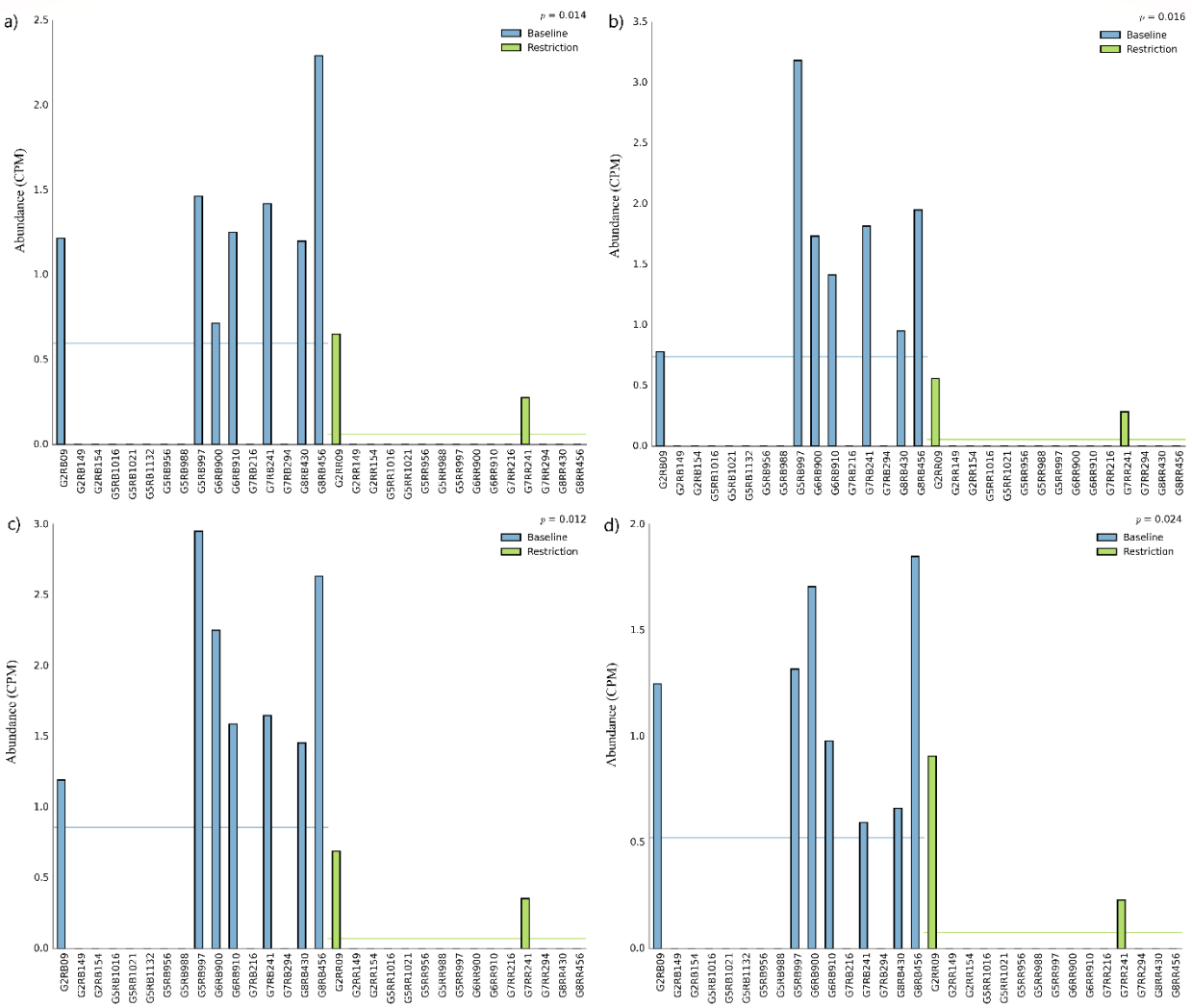


Figure 5.14: A subset of the pathways observed to be significantly ($p<0.05$) low in abundance in restriction rumen samples when compared to baseline; a) L-methionin biosynthesis (p=0.014), b) Pyruvate fermentation to isobutanol (p=0.016), c) Superpathway of branched chain amino acid biosynthesis (p=0.012), d) Superpathway of L-thrionin biosynthesis (p=0.024). Baseline and restriction rumen samples are color

coded in blue and green, respectively. The Y-axis indicates pathway abundance in counts per million (CPM) for each sample and the X-axis indicates samples IDs.

5.3.2.2 Functional dynamics of the fecal microbiome during water restriction

In terms of the fecal microbiome we were able to identify a total of 106 metabolic pathways that had a significant difference ($p<0.05$) in their pathway abundance when between group comparisons were done for baseline and restriction fecal samples. Many biologically important pathways such as (but not limited to) superpathway of branched amino acid biosynthesis, superpathway of aromatic amino acid biosynthesis, glycolysis I (from glucose 6-phosphate), L-histidine biosynthesis, methanogenesis from H_2 and CO_2, formaldehyde assimilation II (RuMP Cycle), amino acid biosynthesis pathways for L-lysine & L-isoleucine, urate biosynthesis, flavin biosynthesis and pyruvate fermentation to isobutanol had significantly low ($p<0.05$) pathway abundances in restriction fecal samples than baseline samples (Figure 5.15).

On the other hand, metabolic pathways such as L-arginine biosynthesis, L-tryptophan biosynthesis, pyruvate fermentation to propanoate I, TCA cycle VIII, D-fructuronate degradation, guanosine nucleotide degradation and superpathway for pyrimidine nucleotides de novo biosynthesis had significantly high ($p<0.05$) pathway abundance in the restriction samples when compared against the baseline samples (Figure 5.16).

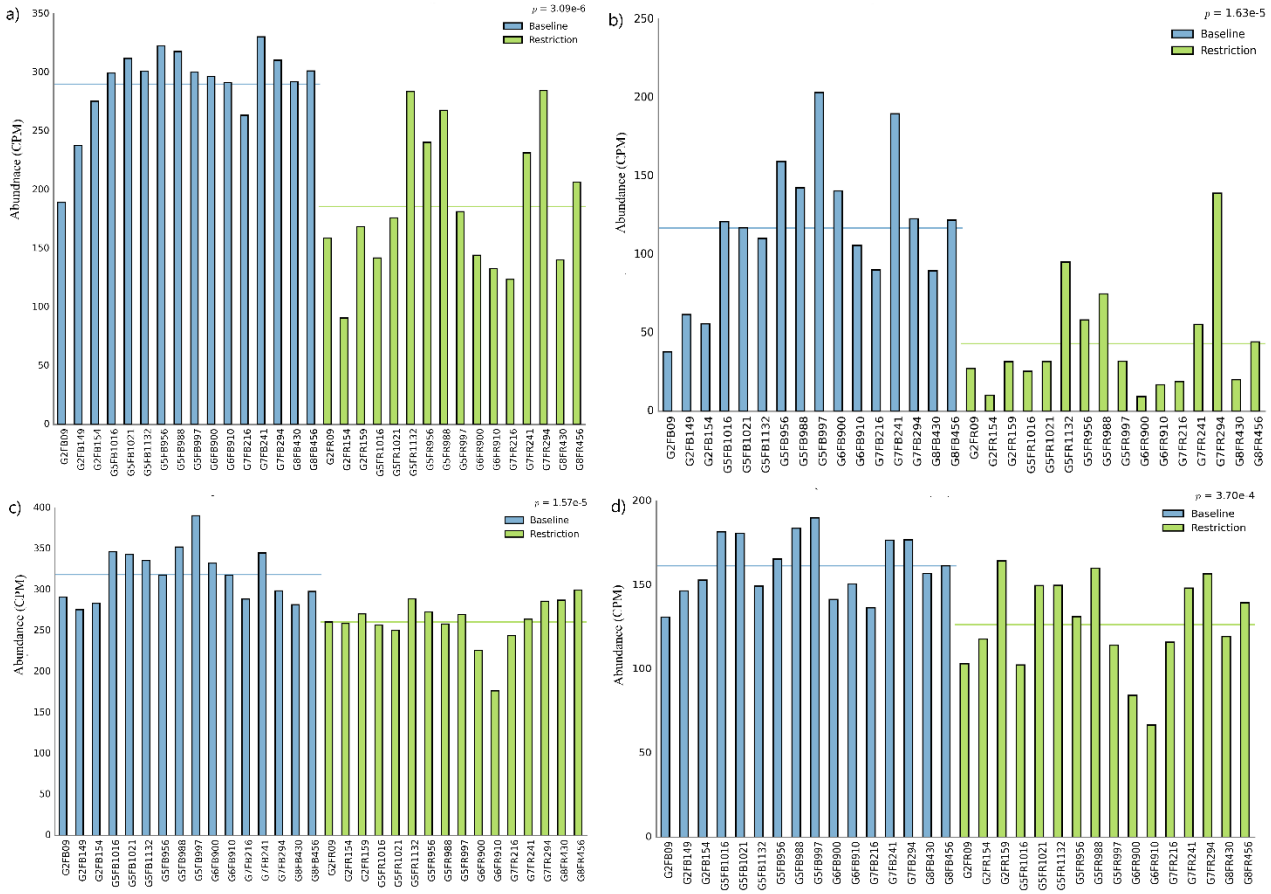


Figure 5.15: A subset of the metabolically important pathways exhibiting significantly low abundance in restriction fecal samples compared to baseline; a) Glycolysis I (from glucose-6-phosphate) (p=3.09e-6), b) Methanogenesis from H_2 and CO_2 (p=1.63e-5), c) Superpathway for branched chain amino acid biosynthesis (p=1.57e-5), d) Superpathway for aromatic amino acid biosynthesis (p=3.7e-4). Baseline samples are indicated in blue and restriction samples are indicated in green. The horizontal lines indicate the mean abundance of that pathway for each group (baseline and restriction). The Y-axis indicates pathway abundance in counts per million (CPM) for each sample and the X-axis indicates samples IDs.

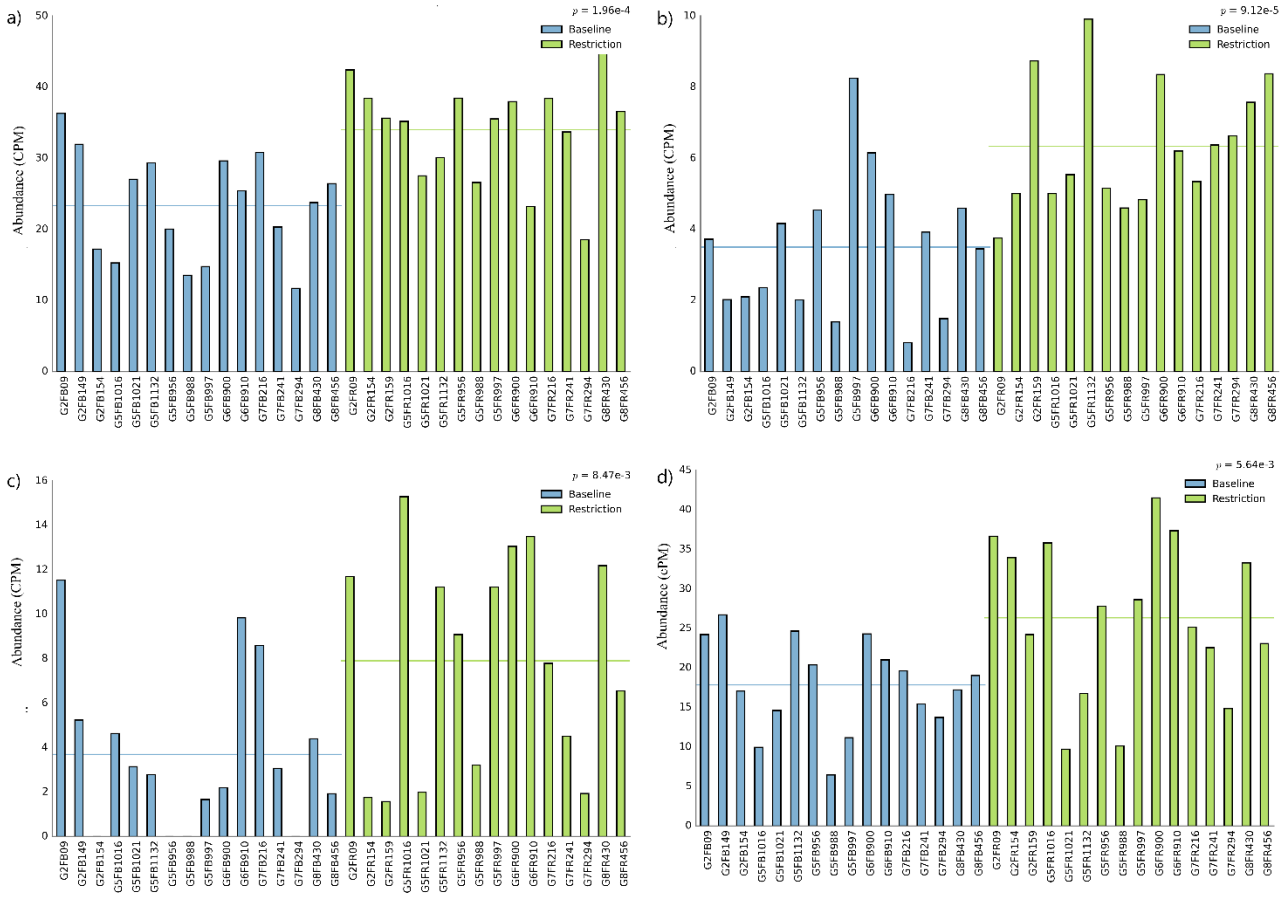


Figure 5.16: Four pathways observed with significantly high abundance in restriction fecal samples compared to baseline; a) D-fructuronate degradation (p=1.96e-4) b) Guanosine nucleotides degradation (p=9.12e-5) c) Superpathway of pyramidine ribonucleotide de novo biosynthesis (p=8.47e-3) d) Pyruvate fermentation to propanoate I (p=5.64e-3). Baseline rumen samples are indicated in blue and restriction rumen samples are indicated in green. Mean pathway abundance for each group is indicated by horizontal lines. The Y-axis indicates pathway abundance in counts per million (CPM) for each sample and the X-axis indicates samples IDs.

5.4 Discussion

The results we obtained during the metagenome sequencing study provides information about the effect of water restriction on the structural and functional dynamics of rumen and fecal microbiomes at the species level. Taxonomic profiling at species level and true functional profiling of a microbial community cannot be achieved only through 16S rRNA gene amplicon sequencing, since only a portion of the 16S rRNA gene is used for taxonomic profiling in these studies. This approach fails to provide reliable taxonomic information beyond genus level and inherently lacks the potential to provide information about the functional capabilities of the microbial community, unless a functional inference software such as PICRUSt (Douglas et al., 2020) or Tax4Fun (Wemheuer et al., 2020) is used. However, the use of these software are sometimes criticized for their predictive nature in generating results, rather than deriving from actual underlying data.

Furthermore, the comparability of 16S rRNA gene amplicon sequencing studies both with its own kind, and metagenome sequencing studies is affected by factors such as the hypervariable region being amplified, the copy number of the 16S rRNA gene, the sequencing technology employed and even on the bioinformatics pipelines used for data analysis (López-García et al., 2018; Winand et al., 2020). The V1 to V9 hypervariable regions are not equally informative across different taxa and differ in their amplification efficiency. Hence, the taxonomy and abundance of the bacterial genera detected can depend on the region choice. The copy number of the 16S rRNA gene vary across different organisms, hence bacterial species with low copy numbers can potentially have

a lesser chance of being detected (Winand et al., 2020). The discrepancies we observed in the predominant taxa identified by the 16S rRNA gene amplicon study and the metagenome study (for both rumen and fecal samples) can be attributed to the reasons mentioned above. Furthermore, since the bioinformatics workflow used for the two studies ultimately depend upon searching against two very differently curated databases (SILVA database and a clade specific marker gene database), it can introduce database bias, thus having a profound effect on the final results.

Non-bacterial microbial communities such as anaerobic fungi and other rumen eukaryotes have proven to be difficult to sequence and characterize due to the considerable technical challenges they present (Seshadri et al., 2018). The vast majority of the fungal diversity described in the rumen and other host-associated or environment-associated microbiomes are based on amplicon sequencing strategies such as ITS (Internally Transcribed Spacer) region sequencing (Lind and Pollard, 2020). Whole metagenome sequencing studies are also used for taxonomic and functional profiling of rumen fungal communities, but are often coupled with multiple parallel enrichment culture experiments in order enrich the fungal community present in the rumen samples (Peng et al., 2021). This is due to the relatively low abundance of rumen fungi compared to total bacterial community, thus contributing a smaller fraction of the reads to shot gun sequence library (Lind and Pollard, 2020). Furthermore, like any other reference database based study, the ability of the current study to detect fungal communities depends on the completeness of the curated databases used for similarity searches. The lack of fungal communities identified during the current study can be mainly attributed to the above

mentioned factors. However, MetaPhlAn was able to successfully detect the fungal communities present in the Zymobiomics microbial community standard used as the positive control during the study. Similar to fungi, the proper detection of rumen viruses through whole metagenome sequencing also requires an efficient viral DNA isolation and enrichment protocol that incorporates membrane filtering to remove other microorganisms such bacteria and protozoa (Namonyo et al., 2018). Such enrichment procedures were beyond the scope of this study, hence the lack of viral communities detected.

The overall species composition data we obtained for both rumen and fecal samples are well supported by existing literature on metagenome sequencing of the bovine rumen and fecal microbiomes. *Sarcina sp* DSM 11001, *Prevotella ruminicola, Streptococcus equinus, Succiniclasticum ruminis,* and *Bacteroidales bacterium* species were categorized as informative taxa in a compendium of 4,941 rumen metagenomes assembled using 283 cattle (Stewart et al., 2019). Furthermore, the authors also found *Methanobrevibacter* and *Methanosphaera* species to constitute 111 out of 126 archaeal genomes they assembled. In another metagenome sequencing study done using cow, sheep, reindeer and red deer rumen samples, *Succiniclasticum ruminis,* and *Sarcina sp* DSM 11001 were identified as species commonly observed in rumen bacteria. They were also part of the 31 rumen uncultured genomes assembled during this study (Glendinning et al., 2021).

Out of the predominant species observed in the fecal microbiome, *Ruminococcaceae bacterium* P7, *Methanosphaera sp* BMS and *Methanobrevibacter millerae* had significant

decreases in their overall abundance during water restriction. This drop in abundance is further reflected in the functional profile as well, where we observed a statistically significant drop in the pathways attributed to these three organisms such as, L-arginine biosynthesis, superpathway of branched chain amino acid biosynthesis, L-isoleucine biosynthesis I, L-lysine biosynthesis III, pyruvate fermentation to isobutanol and L-valine biosynthesis. A significant drop in the methanogenesis pathway abundance was also observed in the restriction fecal samples. The drop in *Methanosphaera sp* BMS and *Methanobrevibacter millerae* may have played a major role in this as well, however methonogenesis pathway was not attributed specifically to any of the two organisms by HUMAnN3.0.

The species level fecal core-microbiome derived using metagenome sequencing data, had notable change during water restriction by losing 3 (out of 5) species it had in the baseline core microbiome, namely; *Ruminococcaceae bacterium* P7, *Methanosphera sp* BMS and *Streptococcus equinus.* This coincides with the drop in the overall relative abundance in these species. The drop in the abundance of *Streptococcus equinus* is also reflected in the functional profile of the fecal microbiome during water restriction. Pathways attributed to S. *equinus* such as sucrose degradation IV (sucrose phosphorylase) pathway and 5-aminoimidazole ribonucleotide biosynthesis I pathway were observed to decrease in abundance during water restriction.

The interactions between the microorganisms that form a microbial community is important in determining microbial community dynamics. Hence, these rich and diverse

relationships are often expected to be reflected in the high throughput microbial community survey datasets. Constructing microbial co-occurrence networks based on their covariation patterns across multiple samples is an effective method of understanding these microbial interactions (Berry and Widder, 2014; Riera and Baldo, 2020). Hence, we used taxonomic profiling data derived from metagenome sequencing to build microbial co-occurrence networks of both rumen and fecal microbiomes under water restriction. We used a novel algorithm called SparCC (Friedman and Alm, 2012) to calculate pairwise correlations between species, since traditional statistical methods such as Pearson and Spearman correlations are known to perform poorly due to the compositional nature of microbiome data, thus giving rise to spurious relationships (Chong et al., 2020).

The co-occurrence network for rumen samples revealed a subnetwork of 14 species that co-vary during water restriction. *Sarcina sp DSM 1101*, *Denitrobacterium detoxificans, Ruminococcus flavefaciens, Streptococcus equinus* and *Methanobravibactor thaueri* could be observed as major hubs of this network and may act as potential keystone species of the rumen co-occurrence network during water restricted conditions. Similarly, for fecal samples we could identify a subnetwork of 17 species held together by *Geobacillus thermodenitrificans, Streptococcus lutetiencis, Streptococcus equinus*, and *Methanobrevibactor millerae* working as hubs and potential keystone species in the fecal co-occurrence network during water restriction.

Kandleria vitulina found to be enriched in the rumen microbiome during water restriction is a homolactic bacteria. Rumen when enriched with homolactic bacteria such as

Kandleria and Sharpea are known to produce less methane, especially in the presence of bacteria such as *Megasphaera spp.* that converts lactic acid to butyrate by producing less H_2 than the traditional carbohydrate fermentation to butyrate. Thus reducing the amount free H_2 available in the rumen that can be utilized for methane production (Kamke et al., 2016; Kumar et al., 2018). The enhanced presence of *Kandleria vitulina* coincides with the increased abundance of homolactic fermentation pathway, observed in the functional profile of the rumen microbiome during water restriction.

Denitrobacterium detoxificans, known to metabolize nitrogen containing compounds such as nitrates in the rumen (Anderson et al., 2016), also emerged as a differentially abundant species in the restriction rumen samples during this study. *D. detoxificans* have been associated with inhibition of methane production in the presence of forage containing nitrate, NPA (3-nitro-1-propionic acid) and NPOH (3-nitro-1-propanol) (Anderson et al., 2016). Hence, water restriction may have an indirect effect towards the reduction of methane production in the rumen via increased abundance of homolactic fermentation, and nitrogen metabolizing bacteria. Thus warranting further investigation on this regard.

Streptococcus infantarius seen differentially abundant in restriction rumen samples is a member of the *S. bovis/S. equinus* complex, a major lactic acid producing bacterium in the rumen (Clarke et al., 2016). Members of this complex are regularly identified as commensals in the GI tract of rumens that play an essential role in proteolysis and carbohydrate degradation. However, they can also become opportunistic pathogens and

cause rumen lactic acidosis primarily due to rapid shifts in diet from forage to concentrate/grain (Jans et al., 2014). No such change in diet occurred during the study and the animals were fed with a constant diet of ~51% wet corn gluten feed, 15% cracked corn, ~28% prairie hay, and ~5% feed additives. However, we did observe a slight but a statistically significant decrease ($p=8.29e-05$) in the average rumen pH of animals during water restriction (from 6.45 to 6.36). Thus, water restriction may have created a favorable environment for the growth of lactic acid producing bacteria like *Streptococcus infantarius,* increasing the susceptibility of cattle to rumen acidosis, during water restricted conditions. However, even though not directly comparable, we could not observe such a trend in the abundance of genus *Streptococcus* during the 16S rRNA gene amplicon sequencing study.

The overall abundance of *Methanosphaera sp BMS, Methanobrevibacter thaueri* and *Methanobrevibacter millerae* decreased in the fecal microbiome during water restriction. They emerged as biomarkers (differentially abundant taxa) for the baseline fecal microbiome as well, meaning that their footprint in restriction samples was reduced significantly. Furthermore, functional profile of the fecal microbiome depicted a significant ($p<0.05$) reduction in the abundance of methanogenesis pathway during water restriction. Other biologically important pathways attributed to these organisms, such as (but not limited to) amino acid biosynthesis (L-lysine, L-isoleucine, L-arginine and L-valine), ribonucleotide biosynthesis and pyruvate fermentation also demonstrated a significant loss in their abundance during water restriction. It is estimated that an approximate 11% of the total methane emitted by a ruminant is produced in the hindgut

(Hook et al., 2010) and most of the protein requirement of the animal is replenished by microbial proteins (Moran, 2005). Therefore, the apparent implications of water restriction on methanogens and microbial amino acid biosynthesis that we have observed during our study, breaks new ground and provides a basis for further investigations to be carried out on this regard.

5.5 References

Anderson, R. C., L. H. Ripley, J. G. P. Bowman, T. R. Callaway, K. J. Genovese, R. C. Beier, R. B. Harvey, and D. J. Nisbet. 2016. Ruminal fermentation of anti-methanogenic nitrate-and nitro-containing forages in vitro. Front. Vet. Sci. 3:1–9. doi:10.3389/fvets.2016.00062.

Berry, D., and S. Widder. 2014. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front. Microbiol. 5:1–14. doi:10.3389/fmicb.2014.00219.

Chen, Yuxin, Yongsheng Chen, C. Shi, Z. Huang, Y. Zhang, S. Li, Y. Li, J. Ye, C. Yu, Z. Li, X. Zhang, J. Wang, H. Yang, L. Fang, and Q. Chen. 2018. SOAPnuke: A MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. Gigascience. 7:1–6. doi:10.1093/gigascience/gix120.

Chong, J., P. Liu, G. Zhou, and J. Xia. 2020. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat. Protoc. 15:799–821. doi:10.1038/s41596-019-0264-1.

Clarke, L. L., R. L. Fathke, S. Sanchez, and J. B. Stanton. 2016. Streptococcus bovis/S. equinus complex septicemia in a group of calves following intramuscular vaccination. J. Vet. Diagnostic Investig. 28:423–428. doi:10.1177/1040638716648364.

Douglas, G. M., V. J. Maffei, J. R. Zaneveld, S. N. Yurgel, J. R. Brown, C. M. Taylor, C. Huttenhower, and M. G. I. Langille. 2020. PICRUSt2 for prediction of metagenome functions. Nat. Biotechnol. 38:685–688. doi:10.1038/s41587-020-0548-6.

Fang, C., H. Zhong, Y. Lin, B. Chen, M. Han, H. Ren, H. Lu, J. M. Luber, M. Xia, W. Li, S. Stein, X. Xu, W. Zhang, R. Drmanac, J. Wang, H. Yang, L. Hammarström, A. D. Kostic, K. Kristiansen, and J. Li. 2018. Assessment of the cPAS-based BGISEQ-500 platform for metagenomic sequencing. Gigascience. 7:1–8. doi:10.1093/gigascience/gix133.

Franzosa, E. A., L. J. McIver, G. Rahnavard, L. R. Thompson, M. Schirmer, G. Weingart, K. S. Lipson, R. Knight, J. G. Caporaso, N. Segata, and C. Huttenhower. 2018. Species-level functional profiling of metagenomes and metatranscriptomes. Nat. Methods. 15:962–968. doi:10.1038/s41592-018-0176-y.

Friedman, J., and E. J. Alm. 2012. Inferring Correlation Networks from Genomic Survey Data. PLoS Comput. Biol. 8:1–11. doi:10.1371/journal.pcbi.1002687.

Glendinning, L., B. Genç, R. J. Wallace, and M. Watson. 2021. Metagenomic analysis of the cow, sheep, reindeer and red deer rumen. Sci. Rep. 11:3–12. doi:10.1038/s41598-021-

81668-9.

Gupta, S., M. S. Mortensen, S. Schjørring, U. Trivedi, G. Vestergaard, J. Stokholm, H. Bisgaard, K. A. Krogfelt, and S. J. Sørensen. 2019. Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. Commun. Biol. 2:1–7. doi:10.1038/s42003-019-0540-1.

Hook, S. E., A. D. G. Wright, and B. W. McBride. 2010. Methanogens: Methane producers of the rumen and mitigation strategies. Archaea. 2010:50–60. doi:10.1155/2010/945785.

Jans, C., L. Meile, C. Lacroix, and M. J. A. Stevens. 2014. Genomics, evolution, and molecular epidemiology of the Streptococcus bovis/Streptococcus equinus complex (SBSEC). Infect. Genet. Evol. 33:419–436. doi:10.1016/j.meegid.2014.09.017.

Kamke, J., S. Kittelmann, P. Soni, Y. Li, M. Tavendale, S. Ganesh, P. H. Janssen, W. Shi, J. Froula, E. M. Rubin, and G. T. Attwood. 2016. Rumen metagenome and metatranscriptome analyses of low methane yield sheep reveals a Sharpea-enriched microbiome characterised by lactic acid formation and utilisation. Microbiome. 4. doi:10.1186/s40168-016-0201-2.

Kumar, S., B. P. Treloar, K. H. Teh, C. M. McKenzie, G. Henderson, G. T. Attwood, S. M. Waters, M. L. Patchett, and P. H. Janssen. 2018. Sharpea and Kandleria are lactic acid producing rumen bacteria that do not change their fermentation products when co-cultured with a methanogen. Anaerobe. 54:31–38. doi:10.1016/j.anaerobe.2018.07.008.

Lind, A. L., and K. S. Pollard. 2020. Accurate and sensitive detection of microbial eukaryotes from metagenomic shotgun sequencing data. bioRxiv. 1–18. doi:10.1101/2020.07.22.216580.

López-García, A., C. Pineda-Quiroga, R. Atxaerandio, A. Pérez, I. Hernández, A. García-Rodríguez, and O. González-Recio. 2018. Comparison of Mothur and QIIME for the Analysis of Rumen Microbiota Composition Based on 16S rRNA Amplicon Sequences. Front. Microbiol. 9:3010. doi:10.3389/FMICB.2018.03010.

Moran, J. 2005. Tropical Dairy Farming. CSIRO Publishing.

Namonyo, S., M. Wagacha, S. Maina, L. Wambua, and M. Agaba. 2018. A metagenomic study of the rumen virome in domestic caprids. Arch. Virol. 163:3415–3419. doi:10.1007/s00705-018-4022-4.

Parks, D. H., G. W. Tyson, P. Hugenholtz, and R. G. Beiko. 2014. STAMP: Statistical analysis of taxonomic and functional profiles. Bioinformatics. 30:3123–3124. doi:10.1093/bioinformatics/btu494.

Peng, X., S. E. Wilken, T. S. Lankiewicz, S. P. Gilmore, J. L. Brown, J. K. Henske, C. L. Swift, A. Salamov, K. Barry, I. V. Grigoriev, M. K. Theodorou, D. L. Valentine, and M. A. O'Malley. 2021. Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. Nat. Microbiol. 6:499–511. doi:10.1038/s41564-020-00861-0.

Rausch, P., M. Rühlemann, B. M. Hermes, S. Doms, T. Dagan, K. Dierking, H. Domin,

S. Fraune, J. Von Frieling, U. Hentschel, F. A. Heinsen, M. Höppner, M. T. Jahn, C. Jaspers, K. A. B. Kissoyan, D. Langfeldt, A. Rehman, T. B. H. Reusch, T. Roeder, R. A. Schmitz, H. Schulenburg, R. Soluch, F. Sommer, E. Stukenbrock, N. Weiland-Bräuer, P. Rosenstiel, A. Franke, T. Bosch, and J. F. Baines. 2019. Comparative analysis of amplicon and metagenomic sequencing methods reveals key features in the evolution of animal metaorganisms. Microbiome. 7:1–19. doi:10.1186/s40168-019-0743-1.

Riera, J. L., and L. Baldo. 2020. Microbial co-occurrence networks of gut microbiota reveal community conservation and diet-associated shifts in cichlid fishes. Anim. Microbiome. 2. doi:10.1186/s42523-020-00054-4.

Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, and C. Huttenhower. 2011. Metagenomic biomarker discovery and explanation. Genome Biol. 12:R60. doi:10.1186/gb-2011-12-6-r60.

Seshadri, R., S. C. Leahy, G. T. Attwood, K. H. Teh, S. C. Lambie, A. L. Cookson, E. A. Eloe-Fadrosh, G. A. Pavlopoulos, M. Hadjithomas, N. J. Varghese, D. Paez-Espino, R. Perry, G. Henderson, C. J. Creevey, N. Terrapon, P. Lapebie, E. Drula, V. Lombard, E. Rubin, N. C. Kyrpides, B. Henrissat, T. Woyke, N. N. Ivanova, W. J. Kelly, N. Palevic, P. H. Janssen, R. S. Ronimus, S. Noel, P. Soni, K. Reilly, T. Atherly, C. Ziemer, A. D. Wright, S. Ishaq, M. Cotta, S. Thompson, K. Crosley, N. McKain, J. J. Wallace, H. J. Flint, J. C. Martin, R. J. Forster, R. J. Gruninger, T. McAllister, R. Gilbert, D. J. Ouwerkerk, A. J. Klieve, R. Al Jassim, S. Denman, C. McSweeney, C. Rosewarne, S. Koike, Y. Kobayashi, M. Mitsumori, T. Shinkai, S. Cravero, and M. Cerón Cucchi. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. Nat. Biotechnol. 36:359–367. doi:10.1038/nbt.4110.

Sharpton, T. J. 2014. An introduction to the analysis of shotgun metagenomic data. Front. Plant Sci. 5:1–14. doi:10.3389/fpls.2014.00209.

Stewart, R. D., M. D. Auffret, A. Warr, A. W. Walker, R. Roehe, and M. Watson. 2019. Compendium of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and enzyme discovery. Nat. Biotechnol. 37:953–961. doi:10.1038/s41587-019-0202-3.

Tickle, T. L., N. Segata, L. Waldron, U. Weingart, and C. Huttenhower. 2013. Two-stage microbial community experimental design. ISME J. 7:2330–2339. doi:10.1038/ismej.2013.139.

Truong, D. T., E. A. Franzosa, T. L. Tickle, M. Scholz, G. Weingart, E. Pasolli, A. Tett, C. Huttenhower, and N. Segata. 2015. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods. 12:902–903. doi:10.1038/nmeth.3589.

Wemheuer, F., J. A. Taylor, R. Daniel, E. Johnston, P. Meinicke, T. Thomas, and B. Wemheuer. 2020. Tax4Fun2: Prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environ. Microbiomes. 15:11. doi:10.1186/s40793-020-00358-7.

Winand, R., B. Bogaerts, S. Hoffman, L. Lefevre, M. Delvoye, J. Van Braekel, Q. Fu, N. H. C. Roosens, S. C. J. De Keersmaecker, and K. Vanneste. 2020. Targeting the 16s

rRNA gene for bacterial identification in complex mixed samples: Comparative evaluation of second (illumina) and third (oxford nanopore technologies) generation sequencing technologies. Int. J. Mol. Sci. 21:1–22. doi:10.3390/ijms21010298.

Appendix

USDA – United States Department of Agriculture

FDA – Food and Drug Administration

USGCRP – United States Global Change Research Program

BW – Body Weight

GRC – Global Rumen Census

ITS1 – Internal Transcribed Spacer 1

NGS – Next Generation Sequencing

RDP – Ribosomal Database Project

WMS – Whole Metagenome Sequencing

VFA – Volatile Fatty Acids

RFI – Residual Feed Intake

PCR – Polymerase Chain Reaction

ddNTP – Dideoxy Nucleotide Triphosphate

HTS – High Throughput Sequencing

SBS – Sequencing by Synthesis

ZMW – Zero-mode Waveguides

ONT – Oxford Nano-pore Technologies

SNP – Single-nucleotide Polymorphism

OTU – Operational Taxonomic Unit

ASV – Amplicon Sequence Variant

DADA2 – Divisive Amplicon De-noising Algorithm 2

LDA – Linear Discriminant Analysis

LEfSe – Linear Discriminant Analysis Effect Size

MAG – Metagenome Assembled Genomes

WSBRC – Willard Sparks Beef Research Center

DMI – Dry Matter Intake

NMDS – Non-parametric Multidimensional Scaling

SRB – Sulfur Reducing Bacteria

ADG – Average Daily Gain

ADFI – Average Daily Feed Intake

RFI – Residual Feed Intake

VITA

HASITHA UPENDRA PREMATHILAKE

Candidate for the Degree of

Doctor of Philosophy

Dissertation: ELUCIDATION OF STRUCTURAL AND FUNCTIONAL CHARACTERISTICS OF THE GUT MICROBIOME OF BEEF CATTLE UNDER WATER STRESS

Major Field: Animal Science

Biographical:

Education:

Completed the requirements for Doctor of Philosophy in Animal Science at Oklahoma State University, Stillwater, Oklahoma in July, 2021.

Completed the requirements for Master of Business Administration at Cardiff Metropolitan University, Cardiff, United Kingdom in October 2014.

Completed the requirements for Bachelor of Science in Bioinformatics at University of Colombo, Colombo, Sri Lanka in November 2011.

Completed the requirements for Bachelor of Science in Information Technology at Sri Lanka Institute of Information Technology, Malambe, Sri Lanka in November 2010.

Experience:
Graduate Research Assistant at Oklahoma State University 2017-2021.
Graduate Teaching Assistant at University of South Dakota 2016-2017.
Graduate Teaching Assistant at University of Colombo 2011-2012.