Identification and Genomic Characterization of Candidate Starch and Lactate Utilizing Bacteria from the Rumen of Beef Cattle

Venkata Bandarupalli
South Dakota State University

Follow this and additional works at: https://openprairie.sdstate.edu/etd

Part of the Animal Sciences Commons

Recommended Citation

This Dissertation - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.
IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE
STARCH AND LACTATE UTILIZING BACTERIA FROM THE RUMEN OF BEEF
CATTLE

BY

VENKATA VINAY KUMAR BANDARUPALLI

A dissertation submitted in partial fulfillment of the requirements for the
Doctor of Philosophy
Major in Animal Science
South Dakota State University
2020
This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Benoit St Pierre
Advisor

Joseph P Cassady
Department Head

Dean, Graduate School
ACKNOWLEDGEMENTS

I am extremely grateful to my advisor Dr. Benoit St-Pierre for his continuous support during my Ph.D. His expertise in the field of microbiology and his invaluable contribution and guidance helped me to become an independent researcher. Without his support and encouragement, the research project would hardly have been completed.

I would like to extend my sincere thanks to my committee members: Dr. Radhey Kaushik, Dr. Joy Scaria, and Dr. George Hamer. They gave their time to provide me with insightful suggestions that helped me to improve my research work.

Many thanks to my fellow labmates for the wonderful times that we spent and for their assistance in my research project. To SDSU Agricultural Experimental Station, thank you for your financial support.

Finally, I deeply thank my parents for their encouragement and guidance that helped me to explore the opportunity of pursuing my Ph.D.
# TABLE OF CONTENTS

**LIST OF FIGURES** ........................................................................................................... vii  
**LIST OF TABLES** .............................................................................................................. x  
**ABSTRACT** ......................................................................................................................... xii  
**CHAPTER 1 .......................................................................................................................... 1  
1.1 Importance of Livestock in Ensuring Human Food Security ................................................... 2  
1.2 Targeting Feed Efficiency for Improved Livestock Production and Sustainability ................. 3  
1.3 Importance of Ruminants for the Sustainable Production of Animal Protein ......................... 3  
1.4 The function of the rumen .................................................................................................. 4  
1.5 The Development of the Rumen ......................................................................................... 5  
2. Methodologies Used in the Analysis of Rumen Microbial Communities ..................................... 7  
  2.1 Culture-Dependent Approaches ......................................................................................... 7  
  2.2 Culture-Independent Approaches ..................................................................................... 8  
   2.2.1 16S rRNA Gene ......................................................................................................... 9  
   2.2.2 Taxonomy Dependent and Independent Methods for 16S rRNA Analysis ................ 11  
  2.3 Omics-Based Approaches ............................................................................................... 12  
   2.3.1 Metagenomics ......................................................................................................... 13  
  2.4 Characterization of the Rumen Microbiome Using the 16S rRNA Gene ............................. 14  
  2.5 Exploring Metabolic Functions of the Rumen Microbiome Using Metagenomics ............. 15  
3.1 Effect of Concentrate or Grain-Based Diets on the Rumen Environment ............................... 17  
3.2 Starch Metabolizing Bacteria in the Rumen Environment ................................................... 18  
3.3 Acidosis, a Digestive Disturbance Resulting from Feeding Concentrate Diets ...................... 20  
3.4 Lactate Metabolizing Bacteria in the Rumen ..................................................................... 22  
4.1 Rationale and Approach .................................................................................................... 24  
4.2 Research Objectives ........................................................................................................ 26  
**References** ......................................................................................................................... 30  
**CHAPTER 2 .......................................................................................................................... 57  
IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE STARCH UTILIZING BACTERIA FROM THE RUMEN OF BEEF CATTLE ............................................. 57  
**ABSTRACT** ......................................................................................................................... 58
1. Introduction ........................................................................................................................................... 59
2. Materials and Methods ................................................................................................................................. 61
   2.1 Sample collection and in vitro rumen culture experiments .......................................................... 61
   2.2 Microbial genomic DNA purification and PCR amplification of the 16S rRNA gene .......... 62
   2.3 Bioinformatic analysis for 16S rRNA gene-based composition analysis ........................................ 63
   2.4 Metagenomic analysis ....................................................................................................................... 65
3. Results ......................................................................................................................................................... 66
   3.1 Comparative analysis of bacterial communities from donor rumen fluid ................................ 66
   3.2 Identification of candidate bacterial starch utilizers from rumen fluid .................................... 67
   3.3 Exploring the metabolic potential of OTUs identified as candidate starch utilizers ............ 68
4. Discussion .................................................................................................................................................... 71
5. Conclusion ................................................................................................................................................ 75
References ..................................................................................................................................................... 90

CHAPTER 3 .................................................................................................................................................. 101
IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE LACTATE UTILIZING BACTERIA FROM THE RUMEN OF BEEF CATTLE ............................................................. 101
ABSTRACT ................................................................................................................................................... 102
1. Introduction ............................................................................................................................................... 103
2. Materials and Methods ............................................................................................................................... 105
   2.1 Sample collection and in vitro rumen culture experiments .......................................................... 105
   2.2 Microbial genomic DNA purification and PCR amplification of the 16S rRNA gene ............ 106
   2.3 Bioinformatic analysis for 16S rRNA gene-based composition analysis ........................................ 107
   2.4 Metagenomic analysis ....................................................................................................................... 108
3. Results ......................................................................................................................................................... 109
   3.1 Comparative analysis of bacterial communities from donor rumen fluid ................................ 109
   3.2 Identification of candidate bacterial lactate utilizers from rumen fluid .................................. 109
   3.3 Exploring the metabolic potential of OTUs identified as candidate lactate utilizers ............ 110
4. Discussion .................................................................................................................................................... 113
5. Conclusion ................................................................................................................................................ 118
References ..................................................................................................................................................... 134

CHAPTER 4 .................................................................................................................................................. 145
FUTURE EXPERIMENTS ................................................................................................................................. 145
4.1 Introduction ........................................................................................................................................... 145
4.2 Research Findings and Future Experiments................................................................. 146
4.3 Research impact............................................................................................................. 148
References........................................................................................................................... 150
LIST OF FIGURES

Figure 1.1. Amylolytic, maltose-fermenting, glucose-fermenting and lactic acid fermenting bacteria involved in the fermentation of high-concentrate diets………27

Figure 2.1. Taxonomic composition of rumen samples used as inocula, prior to setting up batch culture enrichment experiments with starch……………………77

Figure 2.2. RAST gene annotation analysis of contig sets generated from enrichment Trials 2 (blue), 3 (orange) and 4 (grey)……………………………………78

Figure 2.3. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#2 (OTU SD_Bt-00966, affiliated to Prevotella copri). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization……79

Figure 2.4. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU SD_Bt-00010, affiliated to Prevotella albensis). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization…..80

Figure 2.5. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#4 (OTU SD_Bt-00010, affiliated to Prevotella albensis). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization…..81

Figure 3.1. Taxonomic composition of rumen samples used as inocula, prior to setting up batch culture enrichment experiments with lactate………………….120
Figure 3.2. RAST gene annotation analysis of contig sets generated from enrichment Trials 2 (blue) and 3 (orange) .................................................................121

Figure 3.3. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to Butyrivibrio fibrosolvens). A) lactate and pyruvate metabolism ..............................................................122

Figure 3.4. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#2 (OUT LacT2_D14_1-15251, affiliated to Butyrivibrio fibrosolvens). B) butanoate metabolism .................................................................123

Figure 3.5. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to Butyrivibrio fibrosolvens). C) citrate cycle .................................................................124

Figure 3.6. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#3 (OTU LT 3.7-00154, affiliated to Anaerococcus prevotii). A) lactate and pyruvate metabolism .................................................................125

Figure 3.7. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to Anaerococcus prevotii). B) glycolysis pathway .................................................................126
Figure 3.8. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of precited metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*). C) glycogen synthesis and mobilization………………………………………………………………….127

Figure 3.9. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of precited metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*). D) pentose phosphate pathway……………………………………………………………………128
LIST OF TABLES

Table 1.1. PCR and Non-PCR Based Molecular Approaches…………………28

Table 1.2. Overview of amount of starch fermented in the rumen when the livestock are supplemented with different starch sources……………………………………29

Table 2.1. Number of high quality-chimera free sequence reads used for OTU analysis in each trial…………………………………………………………………82

Table 2.2: Shared ruminal bacterial OTUs amongst rumen fluid donors used in this study…………………………………………………………………………………83

Table 2.3. Most abundant# ruminal bacterial OTUs that were shared amongst all rumen fluid donors used in this study…………………………………………………84

Table 2.4. Enrichment of OTUs in response to starch………………………………85

Table 2.5. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU SD_Bt-00966, affiliated to Prevotella albensis)………………………………………………………………………86

Table 2.6. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU SD_Bt-00010, affiliated to Prevotella albensis)………………………………………………………………………87

Table 2.7. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#4 (OTU SD_Bt-00010, affiliated to Prevotella albensis)………………………………………………………………………88
Table 2.8. Number of coding sequences (CDS) assigned to predicted amino acid metabolism-related functions in the contig sets from Trials 2, 3 and 4. 

Table 3.1. Enrichment of OTUs in response to lactate. Abundance is presented as relative abundance (%) of total non-chimeric reads.

Table 3.2. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to *Butyrivibrio fibrisolvens*).

Table 3.3. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*).

Table 3.4. Amino acid metabolizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to *Butyrivibrio fibrisolvens*).

Table 3.5. Amino acid metabolizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*).

Table 3.6. Transporters identified from BLASTp database from two experimental trials.
ABSTRACT

IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE STARCH AND LACTATE UTILIZING BACTERIA FROM THE RUMEN OF BEEF CATTLE

VENKATA VINAY KUMAR BANDARUPALLI

2020

The rumen functions as an anaerobic fermentation chamber where microbial communities ferment feedstuffs into SCFAs. These SCFAs are critical to the ruminant host since they fulfill 70% of its energy requirements. In intensive management systems, ruminants are fed with grain-rich diets, which contain high levels of starch that help meet the energy demands of high producing animals. When grain-based diets are fed to ruminants, ruminal amylolytic and lactate metabolizing bacteria predominate. Grain overload can cause rapid production and accumulation of lactate, resulting in proliferation of amylolytic bacteria and death of lactate metabolizing bacteria. Consequently, high lactate levels result in a decrease in ruminal pH, a condition referred to as acidosis, which can severely impact animal health. A majority of amylolytic and lactate metabolizing bacteria remain uncharacterized, owing to the complexity of the rumen microbiome. In this context, the main objective of the research described in this dissertation was to identify previously uncharacterized bacteria based on their ability to grow in the presence of starch or lactate using an in-vitro batch culture approach. Rumen fluid collected from cannulated beef cows were maintained in laboratory scale bioreactors at a constant physiological temperature (38°C), with three of the replicate cultures supplemented with either starch or lactate. Candidate starch or lactate utilizers
were identified using a 16S rRNA-based analysis by comparing the composition of substrate supplemented and non-supplemented cultures. A total of seven experimental trials were conducted related to starch and lactate, resulting in the identification of ten Operational Taxonomic Units (OTUs) as candidate utilizers. Metagenomics analysis on the samples showing high OTU enrichment in both starch and lactate trials was performed to gain further insights into the functional potential of these OTUs. Gene annotation using RAST revealed enzymes predicted to be associated with metabolic activities of starch and lactate metabolizing bacteria. Therefore, a better understanding of these highly enriched bacteria in starch and lactate trials will provide greater insights for developing probiotics. Successful probiotic development would be beneficial for animal health and performance.
CHAPTER 1

LITERATURE REVIEW
1.1 Importance of Livestock in Ensuring Human Food Security

The global human population is growing at a fast rate, and it is expected to increase from 7 billion, currently, to 9.73 billion by 2050, potentially reaching up to 11.2 billion by 2100 (Godfray et al., 2010) (FAO, 2017). As not only the world population but also urbanization are increasing, concerns arise whether today’s agricultural system will be able to meet the food demands of a growing population (FAO, 2017). According to the Food and Agriculture Organization (FAO), consumption of meat and dairy products is projected to increase by 73% and 58%, respectively, by 2050 (FAO, 2018). Valin et al. (2014) reported that global food production is expected to increase overall by 74% by 2050, with predictions ranging between 54% and 98% depending on the prediction model used (Valin et al., 2014). Deforestation, land degradation, and water scarcity are major factors having a huge negative impact on agricultural productivity, further intensifying the decrease in agricultural productivity globally (FAO, 2017) (FAO, 2011). Production of livestock-derived products, such as milk, meat, and eggs, needs to be increasing rapidly to meet the booming demand of the ever-growing human population (Didanna, 2015) (Bodirsky et al., 2015). Based on models designed to predict global food demand, simply increasing livestock numbers and expanding cropland doesn’t offer a viable solution to meet long-term food demands (Bodirsky et al., 2015). Therefore, one plausible solution for a sustainable global food system is to increase the production efficiency of the livestock sector in order to meet the global demand for animal protein (Council, 2015) (Gerbens-Leenes, Nonhebel, & Krol, 2010).
1.2 Targeting Feed Efficiency for Improved Livestock Production and Sustainability

While global livestock production has been steadily increasing (Byerly, 1967), higher feed cost have had a major impact on the beef, pork, dairy, poultry, and aquaculture industries, because feed cost represent up to three-quarters of total direct costs (Kenny, Fitzsimons, Waters, & McGee, 2018). Accordingly, research on dairy and beef raised in feedlots has prioritized improvement of feed efficiency. This change in strategy is expected to increase profitability for producers and the sustainability of ruminant production systems (Kenny et al., 2018) (Connor, 2015). Over the past two decades, various studies have been conducted to investigate the nutritional requirements of animals to increase feed efficiency (Tona, 2018), leading to the development of modern biotechnological strategies, such as the use of feed additives, crystalline amino acids, probiotics, and prebiotics (Asmare, 2014) (Ravindran, 2013). However, factors such as diet quality and quantity, environmental conditions, host genetics, microbiome composition as well as age can also impact feed efficiency in animals (Soest, 1994) (Bruns, Pritchard, & Boggs, 2005) (Barendse et al., 2007) (Kerr, Kellner, & Shurson, 2015). While research has been conducted for many years on this subject, further progress and improvements are still needed to improve feed efficiency in the livestock industry (Shike, 2013).

1.3 Importance of Ruminants for the Sustainable Production of Animal Protein

Ruminants have the ability to thrive on plant biomass material, which is inedible to humans. This is possible through the synergistic action of complex symbiotic microbial communities that ferment plant fibers into short chain fatty acids (SCFAs) and
produce microbial proteins, which can both be utilized by their host (Kamra, 2005) (Qi et al., 2011). While the host lacks the necessary enzymes for digestion of plant biomass, ruminal microorganisms as a group express a wide range of enzymes that can hydrolyze plant polysaccharides into monomers, which they then convert into ATP to fulfill their energy requirements, and into SCFAs, which are absorbed by the host for use as energy (Yue, Li, & Yu, 2013). Assignment of metabolic functions to the ruminal microbial species that participate in fermentation is thus essential for optimizing rumen function and improving the efficiency of milk and meat production (C. S. Stewart, Fonty, & Gouet, 1988) (Firkins, Karnati, & Yu, 2008) (Eugène, Archimede, & Sauvant, 2004).

However, there are still significant gaps in our knowledge of the key species involved in fermentation activity, and how their metabolism can affect feed efficiency (Kenny et al., 2018).

1.4 The function of the rumen

The rumen is the largest compartment of the foregut of ruminants. It functions as an anerobic fermentation chamber, where microbial symbiotic communities, consisting of bacteria, archaea, protozoa, and fungi, collectively participate in the digestion of feed into SCFAs. These represent the main energy source for the host, providing 60%-70% of its energy requirements. (Russell & Mantovani, 2002; Van Soest, 1994). The role of SCFAs for the host animal has been known for quite some time (Elsden, 1945, Masson and Phillipson, 1951), and it remains a subject of ongoing research of great importance to this day (Cantalapiedra-Hijar, Yanez-Ruiz, Martin-Garcia, & Molina-Alcaide, 2009; Dijkstra, Forbes, & France, 2005; Elsdon, 1945; Flint, 1997; J. Liu et al., 2012; Masson &
Phillipson, 1951; B. Wang et al., 2017). SCFAs are first absorbed through the rumen wall, then they proceed via the bloodstream to the liver and peripheral tissues where they are utilized (den Besten et al., 2013). Acetate can be metabolized for the production of ATP or used as a precursor for the synthesis of fatty acids, while propionate is used as a precursor for gluconeogenesis in the liver. Gluconeogenesis is the main source of glucose for ruminants, which is utilized for specific functions, such as the synthesis of milk lactose, and contributes to overall animal performance such as live weight gain (Balch & Rowland, 1957; Linn, 1988; Urrutia & Harvatine, 2017)

Structural and non-structural carbohydrates represent the primary substrates for SCFA production, with a much greater contribution compared to proteins or lipids. The respective ruminal concentrations for acetate, propionate, and butyrate depend on the type of diet fed to animals. Ruminal acetate concentrations tend to be higher with high fiber diets, whereas propionate is in higher abundance when concentrate-based diets are fed. Other factors that can affect SCFA production and their respective abundance include feed intake, frequency of feeding, as well as composition of the ruminal microbiome (Dijkstra et al., 2005).

1.5 The Development of the Rumen

Ruminants are born with an underdeveloped rumen, and their digestion system functions like that of monogastric animals until they transition to a solid-based diet. Suckling triggers a reflex mechanism that causes the formation of the esophageal groove, an anatomical structure that allows ingested milk to bypass the rumen and enter the abomasum directly, which is the true stomach of ruminants, where milk is digested by
host enzymes (Church, 1988; Davis, 1988; Van Soest, 1994). The transition from pre-
ruminant to ruminant is one of the most important stages in the life of a young ruminant
(Jiao et al., 2015). Indeed, the change from suckling to a solid diet requires the
development of microbial communities that can metabolize plant biomass as well as
anatomical structures for absorption of microbially produced SCFAs. These two
processes typically proceed together, and their outcome can greatly impact the future
performance of the animal, particularly in intense production systems.

The development of ruminal microbial communities from a pre-ruminant state to
fully functional involves multiple cycles of microbial succession, where the composition
of communities changes in response to changes in diet and in coordination with host
development. Microbial colonization of the rumen starts immediately after birth, from
exposure to a wide array of microorganisms present in the immediate environment, such
as the dam’s udder, milk, saliva, and manure, as well as from other animals (Church,
1988). While aerobic and facultative anaerobic taxa predominate after birth, they are
gradually replaced over a period of 9-13 weeks by taxa that are exclusively anaerobic
(Beharka, Nagaraja, Morrill, Kennedy, & Klemm, 1998; Bryant, Small, Bouma, &
Robinson, 1958; Fonty, Gouet, Jouany, & Senaud, 1987; Minato, Otsuka, Shirasaka,
Itabashi, & Mitsumori, 1992). Functional groups appear to be established early, with
amylolytic bacteria appearing first in 1-day-old calves, while cellulolytic bacteria appear
later, after 3-5-days (Bryant, Small, Bouma, & Robinson, 1958; Fonty et al., 1987;
Minato et al., 1992). Using the 16S rRNA gene as a marker for bacterial composition,
Jami et al. (Jami, Israel, Kotser, & Mizrahi, 2013) investigated the composition of
ruminal bacterial communities in four different age groups: 1–3 day-old calves, 2 month-
old calves, 6 month-old heifers, and 2 year-old lactating dairy cows. While *Firmicutes*, *Bacteroidetes*, and *Proteobacteria* were the main phyla across all age groups, *Firmicutes* were found to be the most abundant in the youngest animals (1–3 day), in contrast to the older group where *Bacteroidetes* were the most dominant. Li et al. (2012) also identified *Firmicutes, Bacteroidetes*, and *Proteobacteria* as the most abundant groups in growing bull calves, with *Firmicutes* being the most abundant at the 14-day stage, while *Bacteroidetes* were predominant in older animals later (42 days and 12-months). This study also indicated that metabolic potential developed as early as the pre-ruminant stage, since enzymes predicted to be involved in metabolizing carbohydrates were identified. Further investigations of ruminal microbial community transitions will be required to better define the impact of solid diets and their ingredient composition on this process as animals develop and mature (Jami et al., 2013; R. W. Li, Connor, Li, Baldwin Vi, & Sparks, 2012; Rey et al., 2014).

2. Methodologies Used in the Analysis of Rumen Microbial Communities

Due to their nature, ruminal microbial communities are challenging to investigate. In the following sections, various strategies and approaches that have been developed over the years for this purpose will be described.

2.1 Culture-Dependent Approaches

The main contribution of cultivation-based methods is to provide a deeper understanding of the metabolic potential of individual isolates and their specific
contribution to rumen function. For many years, these were the main approaches used for investigating ruminal microbial species, with a primary focus on the ones responsible for metabolizing plant biomass substrates that are resistant to host digestive enzymes (Ze, Le Mougen, Duncan, Louis, & Flint, 2013). A number of bacterial species were characterized as keystone species, based on their role in initiating the digestion of feed substrates. For instance, *Ruminococcus flavefaciens*, *Ruminococcus albus*, and *Fibrobacter succinogenes*, which are involved in metabolizing cellulose, were identified and characterized using traditional culturing methods (Shi, Odt, & Weimer, 1997). Ruminal hemicellulolytic, amylolytic, and lactate utilizing species were also described using similar approaches (Dehority & Grubb, 1977). According to a recent survey, at least 146 ruminal bacterial species have so far been cultured, and these are distributed across 88 different genera and seven phyla (Creevey, Kelly, Henderson, & Leahy, 2014). *Firmicutes* is the phylum with the highest number of cultured ruminal species up to this point, which include forty-five species that are affiliated to *Lachnospiraceae* (Creevey et al., 2014).

### 2.2 Culture-Independent Approaches

While growth of isolates as individual cultures remains a gold standard when characterizing the metabolic activities of microbial symbionts, one limitation of culture-based methods is their inability to provide community-level assessments, such as composition, diversity, richness, and evenness. Molecular methods, primarily based on analysis of DNA sequences, have provided an effective means of determining the composition of microbial communities without requiring culturing of isolates (Agrawal,
Agrawal, & Shrivastava, 2015; Medlin & Orozco, 2017). The most popular DNA-based strategy has been the use of the small ribosomal RNA subunit (16S rRNA), which has, since its implementation (Pace, Stahl, Lane, & Olsen, 1986), facilitated the investigation of a very wide range of microbial ecosystems and has revealed a level of bacterial species diversity and richness far greater than what had been previously estimated (Hamady & Knight, 2009).

2.2.1 16S rRNA Gene

The small ribosomal RNA subunit (16S rRNA) is the most commonly used phylogenetic marker for prokaryotic species. Its ubiquity, genetic stability, and length are some of the characteristics that have made this gene a “gold standard” for characterization of unknown microbial species (Ibal, Pham, Park, & Shin, 2019) (Větrovský & Baldrian, 2013). It is an optimal marker when working with ecological habitats that harbor poorly characterized bacterial communities whose members are difficult to culture, such as communities living in symbiosis with a host (Arumugam et al., 2011; Costello et al., 2009; Turnbaugh et al., 2009) or in a physical environment such as in soils or oceans (Gilbert et al., 2009; Hackl, Zechmeister-Boltenstern, Bodrossy, & Sessitsch, 2004; Woo, Lau, Teng, Tse, & Yuen, 2008). In clinical microbiology, 16S rRNA gene sequencing has proven to be an effective and comprehensive way of providing information about poorly characterized pathogenic bacteria, thus contributing to a better understanding of diseases and selection of appropriate treatment (Clarridge, 2004).
The 16S rRNA gene is approximately 1,550 bp long, and its basic structure consists of alternating conserved and hypervariable regions. The conserved regions can be used as targets for PCR primer design, while variable regions can be used for classification at various taxonomic levels, which can range from phylum to strain level identification (J. B. Patel, 2001). Nine distinct hypervariable regions (V1-V9), each differing in nucleotide length, sequence, and phylogenetic resolution, can be used for taxonomic classification (Chakravorty, Helb, Burday, Connell, & Alland, 2007; Youssef et al., 2009) (Yu & Morrison, 2004) (Z. Liu, Lozupone, Hamady, Bushman, & Knight, 2007) (Zongzhi Liu, DeSantis, Andersen, & Knight, 2008; Q. Wang, Garrity, Tiedje, & Cole, 2007).

During the period when Sanger sequencing was the only available option, 16S rRNA-based approaches were limited by clone library construction and costs. The development of Next Generation Sequencing platforms revolutionized this approach by providing much higher sequence yields and greater species coverage for complex microbial communities (Quail et al., 2012). However, most Next Generation Sequencing platforms are limited in the length of individual sequence reads that they can generate, and are thus unable to provide the full sequence of individual 16S rRNA genes. To counter this limitation, a preferred strategy to target a sub-region of the 16S rRNA for sequencing instead of the full-length gene was developed. In an effort to identify sub-regions of the 16S rRNA genes that would be optimal for this purpose, Kim and Morrison (2011a) used an *in-silico* approach to compare the phylogenetic resolution of different hypervariable regions to their corresponding full length 16S rRNA gene sequences. They determined that V1-V3 and V1-V4 provided the best phylogenetic resolution.
While the use of the 16S rRNA gene for composition analysis of complex microbial communities has become very popular because it provides many advantages (Chaucheyras-Durand & Ossa, 2014), there are a number of drawbacks to its use. The variation in gene copy numbers amongst different bacterial species, which can range from 1 to 15 copies, can affect estimates of relative abundance, thus impacting bacterial composition analyses (Klappenbach, Saxman, Cole, & Schmidt, 2001) (Case et al., 2007). As 16S rRNA-based analyses typically use amplification by PCR to generate DNA templates for sequencing, this process can introduce biases caused by the primers used, amplification efficiency and/or the production of artifacts (Chandler, 1998; Costea et al., 2017; de Liptihay, Enzinger, Johnsen, Aamand, & Sørensen, 2004; Forney, Zhou, & Brown, 2004; Morgan, Darling, & Eisen, 2010; Polz & Cavanaugh, 1998; Sipos et al., 2007; Webster, Newberry, Fry, & Weightman, 2003).

2.2.2 Taxonomy Dependent and Independent Methods for 16S rRNA Analysis

Taxonomy dependent methods, also known as supervised methods, rely on information from reference databases to assign sequences from experimental datasets to taxonomic groups. The major limitation for this approach is its dependence on accurate and robust databases, which can result in its inability to assign bacterial species from poorly characterized phylogenetic lineages to available taxonomic groups (Sedlar, Kupkova, & Provaznik, 2017). The k-neighbor based tool RDP classifier is one of the most widely used implementations of this approach. Its associated database, the Ribosomal Database Project (RDP), has a reported ~3 million reference 16S rRNA sequences, which include 44% of its bacterial reference sequences and 15.3% of its
archaeal references with a length of at least 1,200 bp (Cole et al., 2014). As research efforts continue to be invested in microbial habitats and further advancements in sequencing technologies, 16S rRNA gene sequence databases are expected to continue to expand and become increasingly more comprehensive.

In contrast, taxonomy independent methods, also known as unsupervised methods, bypass the need for reference database comparisons by using a clustering-based approach. In this case, sequences in an experimental dataset are compared to each other, then divided into distinct groups according to a pre-determined threshold in nucleotide differences, commonly referred to as genetic distance. This approach is an effective strategy for higher resolution when exploring complex microbial systems that consist mostly of uncharacterized bacterial species, such as would be the case for the rumen microbiome for instance (Kim, Morrison, & Yu, 2011a).

2.3 Omics-Based Approaches

While DNA-based approaches have originally relied heavily on PCR-based strategies, they have gradually been overshadowed by the development of high throughput shotgun sequencing technologies, including metagenomics and metatranscriptomics. These have increased our ability to understand microbial community composition and functional potential at greater depths (Franzosa et al., 2015) (Segal et al., 2019). (Table 1.1).
2.3.1 Metagenomics

One of the first reports on the implementation of metagenomics described a collective analysis of microbial genomes from soil microorganisms (Handelsman et al. 1998), which bypassed the need for traditional culture-dependent methods. As a DNA-based method, metagenomics require extraction of microbial genomic DNA from habitats of interest, just like 16S rRNA based methods, but it does not require PCR, since genomic DNA is sequenced directly in a random fashion, a process commonly designated as ‘shotgun sequencing’. When investigated microbial communities consist of poorly characterized bacterial species for which reference genomes are not available, sequence datasets, which consist of high numbers of short DNA sequences, are analyzed primarily through identification of potential coding sequences and gene annotation. For this purpose, sequence reads can be assembled into contigs, which correspond to segments of microbial genomic DNA from bacterial species of the sampled microbial habitat. A number of online and standalone software packages have been developed for this purpose, including BBAP (Lin et al., 2017), Genovo (Laserson, Jojic, & Koller, 2011; Sato & Sakakibara, 2013), IDBA-UD (Peng, Leung, Yiu, & Chin, 2012), IVA (Hunt et al., 2015), MAP (Lai, Ding, Li, Duan, & Zhu, 2012), MegaGTA (Medlin & Orozco, 2017), MEGAHIT (D. Li, Liu, Luo, Sadakane, & Lam, 2015; D. Li et al., 2016), MetaVelvet (Namiki, Hachiya, Tanaka, & Sakakibara, 2012), Omega (Haider et al., 2014), PRICE (Ruby, Bellare, & DeRisi, 2013), and Ray Meta (Boisvert, Raymond, Godzaridis, Laviolette, & Corbeil, 2012). Once contigs have been constructed, gene annotation tools such as RAST (Aziz et al., 2008), MEGAN (Huson & Weber, 2013), IMG/M (Markowitz et al., 2012), Metarep (Goll et al., 2010), or MicroScope (Vallenet et
al., 2013), can be used for identification of potential coding sequences in order to match them to proteins of known function. Alternatively, gene annotation can be performed on individual reads without contig assembly, using tools such as MG-RAST (Meyer et al., 2008). These tools are key in elucidating metabolic pathways of microbial communities from any habitat or environment, by providing critical information on genes present in these communities (Deusch, Tilocca, Camarinha-Silva, & Seifert, 2015).

2.4 Characterization of the Rumen Microbiome Using the 16S rRNA Gene

Kim et al. (2011b) examined bacterial diversity in the rumen by conducting a meta-analysis of curated 16S rRNA gene sequences of rumen origin. They have used 13,478 bacterial and 3516 archaeal sequences to examine the bacterial diversity from the RDP database. From this analysis, at least 19 bacterial phyla have so far been identified in the rumen, with *Firmicutes*, *Bacteroidetes*, and *Proteobacteria* being the most dominant groups, accounting for 57.8%, 26.7%, and 6.9% of the total bacterial sequences (Kim, Morrison, & Yu, 2011b). In *Firmicutes*, 90.6% of sequences are assigned to class Clostridia, while in *Bacteroidetes*, 88.5% of the sequences are assigned to class Bacteroidia. The predominant genera in phylum *Firmicutes* include *Butyrivibrio*, *Acetivibrio*, *Ruminococcus*, *Succinlasticum*, *Pseudobutyribrio*, and *Mogibacterium*, whereas *Prevotella* is the most dominant genus in *Bacteroidetes* accounting for 41.5% of the *Bacteriodetes* sequences (Kim et al., 2011b).
2.5 Exploring Metabolic Functions of the Rumen Microbiome Using Metagenomics

Diet is a key factor modulating ruminal microbial community structure, likely because the various enzyme activities that are required for metabolizing feed stuffs are expressed by specialized ruminal microorganisms (Carberry, Kenny, Han, McCabe, & Waters, 2012). Linking microbial species to their preferred substrates represents critical information to further understand how rumen function impacts animal health and performance (Deusch et al., 2015; Morgavi, Kelly, Janssen, & Attwood, 2013). While 16S rRNA gene-based analyses provide a comprehensive way of identifying unknown bacterial and archaeal species, its major limitation is its inability to accurately provide information on their respective functions and metabolic capabilities (Case et al., 2007) (von Mering et al., 2007). It is for this reason that intensive metagenomics research on the rumen microbiome has been ongoing.

Metagenomic studies of the rumen microbiome tend to focus on identifying candidate enzymes involved in structural polysaccharide hydrolysis (Brulc et al., 2009; D. D. Patel et al., 2014; Lingling Wang, Hatem, Catalyurek, Morrison, & Yu, 2013). As one of the earlier metagenomic reports, Ferrer et al. (2005) identified 12 different esterases, nine different forms of endo-β-1,4-glucanases, and one cyclodextrin enzyme from a dairy cow fed a forage-based diet of ryegrass and clover (Ferrer et al., 2005). Hess et al. (2011) later identified 57 different types of cellulose hydrolyzing enzymes out of 90 expressed proteins from the bovine rumen using metagenomic DNA libraries (Hess et al., 2011). More recently, metagenomics analyses on rumen digesta samples from Holstein–Friesian crossbred steers that had been fed finger millet straw revealed a diverse
carbohydrate-metabolizing enzyme profile, with glycoside hydrolases being the most abundant class of enzymes (Jose, Appoothy, More, & Arun, 2017). Wang et al. (2019) reported on a study that investigated the carbohydrate-metabolizing enzymes diversity of Holstein cows that had been fed diets with either high forage (70%) or low forage (30%) diets. Glycoside hydrolases, which are predicted to be involved in the digestion of fiber, were found to be the most abundant class of enzymes identified in both diet groups (Lijun Wang, Zhang, Xu, Xin, & Zhang, 2019). To date, the most comprehensive study on the rumen microbiome using metagenomics reported the assembly of 913 draft genomes from datasets generated from 43 Scottish cattle (R. D. Stewart et al., 2018). Using metagenomic binning and Hi-C-based proximity-guided assembly, the contigs as a group encoded for 1,979,391 genes, of which 69,678 coded for enzymes predicted to be involved in carbohydrate utilization (R. D. Stewart et al., 2018).

In contrast, only a limited number of studies have reported the presence of starch hydrolyzing or lactate utilizing enzymes in the rumen environment using metagenomics. Pitta et al. (2016) found high oligosaccharide enzyme activity, as well as α-amylases, in primiparous and multiparous cows that had been fed high concentration diets at different stages of lactation, an observation consistent with the abundance of starch in the experimental diets (Pitta et al., 2016). Metagenomics analyses can also reveal genes or coding sequences involved in other bioactivities, such as novel lipases (Bayer, Kunert, Ballschmiter, & Greiner-Stoeffele, 2010; K. Liu et al., 2009), neutralization of plant toxins (Morgavi et al., 2013), bacterial antibiotic resistance genes (Thomas et al., 2017), and synthesis of specific compounds such as conjugated linoleic acids during milk-fat depression (Morgavi et al., 2013).
While metagenomics has become an indispensable tool in the field of rumen microbiology towards elucidating and understanding the functions of uncultured bacteria and their impact on host nutrition, this approach is still in its infancy (Singh, Gautam, Verma, Kumar, & Singh, 2008). Since functional annotation is performed primarily for unknown or poorly characterized microorganisms, one of the main challenges has remained the limited number of available reference genomes for ruminal microorganisms (Denman & McSweeney, 2015).

3.1 Effect of Concentrate or Grain-Based Diets on the Rumen Environment

Grains are rich energy sources because they contain high levels of starch, a glucose polymer that is more readily hydrolyzed than structural polysaccharides such as cellulose and hemicellulose. Grains are used in intensive animal production systems to maximize growth potential, as they provide elevated levels of starch that can satisfy periods of high-energy demands by animals. Amongst grains typically available for livestock production, wheat contains the highest amount of starch (77%), followed by corn and sorghum (72%), then barley and oats (57-58%) (Ferraretto, 2017; Seoane, 1992). However, the amount of starch is not the only factor to consider, as the availability or digestibility of starch in the rumen is not necessarily proportional to its abundance in a particular source (Table 1.2). In addition to the type of grain, other factors can affect starch digestion rates in the rumen, including the processing method used, amount fed, other ingredients present in the diet, intake or amount of feed ingested by the animals, feeding management, and the ability of ruminal microbial communities to adapt their
metabolic activities to elevated amounts of starch (Hoffman et al., 2011; Mills, France, & Dijkstra, 1999).

3.2 Starch Metabolizing Bacteria in the Rumen Environment

The transition from forages to a highly fermentable carbohydrate (concentrate) diet alters the rumen microbiome, with starch hydrolyzing bacteria becoming more predominant compared to fibrolytic bacteria. Robert E. Hungate, pioneer in rumen microbiology, found alterations in the rumen microbiome when animal diets were supplemented with high levels of grain, with significant reductions in cellulolytic bacterial and protozoal counts, while gram-positive bacteria, particularly *Streptococcus bovis*, increased in abundance (Hungate, Dougherty, Bryant, & Cello, 1952). Similarly, starch metabolizing bacteria were found to range in relative abundance between 90% and 95% in animals fed grain-based diets (Leedle, Bryant, & Hespell, 1982). Bryant et al. (1958a) reported *Succinimonas amylolytica* as a starch utilizing ruminal species, and showed that it produced succinate as a main SCFA end product (Bryant, Small, Bouma, & Chu, 1958). Mackie and Gilchrist (1979) identified additional species belonging to the genera *Butyrivibrio*, *Eubacterium*, and *Lactobacillus* as major players in starch fermentation in the rumen of sheep. In their study, Cotta et al. (1988) tested ruminal starch-utilizing bacterial species *in vitro* by culturing on starch-based medium, and reported that *Streptococcus bovis JB1*, *Bacteroides ruminicola 23* and *B14*, and *Butyrivibrio fibrisolvens A38* and 49 were found to have the highest starch hydrolyzing activity. In a later study, McAllister et al. (1990) determined the starch fermentation activity of a few selected species on different diets that included barley, maize, or wheat,
and identified *Butyrivibrio fibrisolvens A38* as a species capable of high amylolytic activity in their assay (McAllister, Cheng, Rode, & Forsberg, 1990). *Ruminobacter amylophilus*, an obligate anaerobe expressing amylase, amylopectinase, and pullulanase enzymes, has also been described as an excellent model to investigate starch fermentation in the rumen (Anderson, 1995). More recently, other amylolytic bacteria have been isolated from the rumen (Figure 1.1) (Nagaraja & Titgemeyer, 2007).

Among the various starch-utilizing species characterized from the rumen, *Streptococcus bovis* has been studied more extensively because it is considered to be an important starch-fermenter (P. N. Hobson & Macpherson, 1952; Walker, 1965; Walker & Hope, 1964). The cell density of this species when animals are on forage-based diets is only 10^4-10^7 cells/grams, while its cellular density can increase up to 10^{11} when grain-based diets are used (Nagaraja & Titgemeyer, 2007). High growth rates for *S. bovis* can be observed during sudden transitions from forage-based diets to concentrate-diets (Wells, Krause, Callaway, & Russell, 1997). Using a qPCR approach, *Prevotella bryantii*, *Selenomonas ruminantium*, and *Mitsuokella multiacidus* were found to increase in Holstein cows that were switched from a basal diet to a high-grain diet (K. Tajima et al., 2001). Using the same approach, *Streptococcus bovis*, *Selenomonas ruminantium*, and *Prevotella bryantii* were similarly confirmed to increase in abundance during adaptation to high-concentrate diets (Fernando et al., 2010).

While a number of starch utilizers, such as members of the genera *Butyrivibrio*, *Eubacterium*, *Lactobacillus*, *Prevotella*, *Ruminobacter*, *Selenomonas*, and *Streptococcus*, have been isolated using traditional-culturing approaches (Peter N Hobson & Stewart, 2012; Kiyoshi Tajima et al., 2000), culture-independent studies have revealed a much
greater diversity and species richness in animals fed with concentrate-based diets than what had been estimated. Using Terminal Restriction Fragment Length Polymorphism (T-RFLP), Fernando et al. (2010) reported that Bacteroidetes, primarily members of the genus Prevotella, were more dominant in animals fed a high-concentrate diet compared to animals fed on prairie hay. Similarly, but using PCR-DGGE, greater ruminal bacterial diversity was reported in animals fed a high concentrate diet without forage than in animals fed a standard high concentrate diet (R. M. Petri, Forster, Yang, McKinnon, & McAllister, 2012). In their study, Zhang et al. (2017) found that the relative abundance of Fibrobacter decreased significantly in the rumen of Holstein heifers when they were fed increasing levels of concentrate, whereas members of the genera Christensenella and Turicibacter increased in concentrate-based diets (Zhang et al., 2017).

3.3 Acidosis, a Digestive Disturbance Resulting from Feeding Concentrate Diets

A sudden change in diet from high forage to inclusion of high levels of starch will cause rapid proliferation of starch utilizing bacteria and a faster rate of SCFAs production, resulting in a decrease in ruminal pH. When ruminal pH remains low for extensive periods, a condition referred to as acidosis, an animal may experience health problems. One of the causes of this condition is the accumulation of lactate, a ruminal SCFAs that is more likely produced from amylytic bacteria than from cellulolytic bacteria, and which can further exacerbate the impact of accumulating main SCFAs on rumen pH. Typically, bacterial species that predominant under acidotic conditions have the ability to continue to metabolize starch despite the accumulation of lactate, their primary end product. These include Selenomonas ruminantium (Caldwell & Bryant,
1966; Latham, Sharpe, & Sutton, 1971; Ricke, Martin, & Nisbet, 1996) *Mitsuokella multiacidus* (Peter N Hobson & Stewart, 2012; Paster, Dewhirst, Olsen, Fraser, & Socransky, 1995), as well as *Lactobacillus ruminis* and *Lactobacillus vitulinus* (Sharpe, Latham, Garvie, Zirngibl, & Kandler, 1973). Many others are likely involved, as Krogh et al. (1963) reported the isolation of 117 *Lactobacilli* strains from cattle and sheep during acidotic condition (Krogh, 1963).

A number of studies on acidosis have shown that diversity, richness, and evenness of bacterial communities are reduced significantly, which likely impact the production efficiency of affected animal (Fernando et al., 2010; Khafipour, Li, Plaizier, & Krause, 2009; S. Li, Khafipour, Yoon, Plaizier, & Scott, 2016; S. Y. Mao, Zhang, Wang, & Zhu, 2013b; R. M. Petri et al., 2012). At the phylum level, most studies have reported an increase in the relative abundance of *Firmicutes* with a corresponding decrease in *Bacteroidetes* under these conditions (R. W. Li et al., 2012; S. Mao, 2017; S. Y. Mao, Zhang, Wang, & Zhu, 2013a; Renee M. Petri et al., 2013; J. C. Plaizier et al., 2017), with only a limited number of reports indicating an increase in *Bacteroidetes* (Fernando et al., 2010). More variation amongst studies is seen when the composition of bacterial communities is analyzed at higher taxonomic resolution. For instance, an increase in *Ruminococcus, Atopobium, Bifidobacterium*, and unclassified *Clostridiales* was observed in SARA induced animals, while members of the genera *Prevotella*, *Treponema*, *Anaeroplasma*, *Papillibacter*, and *Acinetobacter* were found in lower abundance (S. Y. Mao et al., 2013a). In their study, Petri et al. (2013b) reported higher abundances of *Prevotella, Acetitomaculum, Pseudobutyrivibrio*, and *Selenomonas* in clinically acidotic heifers (Renee M. Petri et al., 2013), while McCann et al. (2016) found increased relative
abundances of *Prevotella*, *Ruminococcus*, *Streptococcus*, and *Lactobacillus* in SARA challenged animals (McCann et al., 2016). Plaizier et al. (2017b) reported higher abundance of *Sharpea*, *Ruminococcus*, *Shuttleworthia*, and *Megasphaera* in SARA challenged animals, while levels of *Anaerostipes* and of the candidate taxon *CF231* decreased (Plaizier, Li, Tun, & Khaipour, 2017). Higher abundance of *Ruminococcus* was observed in SARA challenged animals by Nagata et al. (2018), while *Prevotella*, *Eubacterium*, and *Oscillibacter* were found in lower abundance (Nagata et al., 2018). Finally, a metatranscriptomic analysis showed that the relative abundance of *Cutibacterium avidum*, *Cutibacterium granulosum*, *Bacteroides vulgatus*, and *Dermacoccus nishinomiyensis* increased in response to SARA, while *Prevotella denticola* and *Prevotella scopos* showed a decrease in their relative abundance (Ogunade, Pech-Cervantes, & Schweickart, 2019).

### 3.4 Lactate Metabolizing Bacteria in the Rumen

Lactate is a ruminal SCFA that accumulates to higher concentrations when animals are fed with highly fermentable carbohydrate rich diets. Since high levels of lactate can be detrimental to the health of ruminants, lactate-utilizing bacterial species have generated a great deal of interest as key players to control lactate accumulation in the rumen. *Megasphaera elsdenii* and *Selenomonas ruminantium* subsp. *lactilytica* have been the best described lactate utilizers in animals adapted to grain diets (Huber, Cooley, Goetsch, & Das, 1976; Mackie, Gilchrist, Robberts, Hannah, & Schwartz, 1978), with the former considered to be the key player in lactate fermentation because of its tolerance to acidity (Therion, Kistner, & Kornelius, 1982). Indeed, the abundance of *Megasphaera*
*Megasphaera elsdenii* has been reported to increase during the adaptation to high concentrate diets and to remain elevated when these diets are maintained (Fernando et al., 2010) (R. M. Petri et al., 2012). In their study, Counotte et al. (1981) found that *Megasphaera elsdenii* was responsible for metabolizing up to 74% of ruminal lactate. Notably, *Megasphaera elsdenii* can produce butyrate or propionate when metabolizing lactate, through expression of enzymes such as NAD-independent D-lactate dehydrogenase and lactate racemase (Hino & Kuroda, 1993). Intriguingly, both of these enzymes that are needed for lactate utilization are suppressed when *Megasphaera elsdenii* is grown on glucose (Hino & Kuroda, 1993). By comparison, *Selenomonas ruminantium* has a very slow growth rate on lactate medium, and can only utilize D-lactate as this strain has a NAD-independent D-lactate dehydrogenase (iD-LDH) enzyme (Asanuma & Hino, 2005; Russell & Baldwin, 1978).

Amongst the limited number of other known ruminal lactate utilizers, five strains of *Propionibacterium acnes* were originally identified by growth on lactate-supplemented (Gutierrez, 1953), and *Fusobacterium necrophorum*, a pathogenic bacterium responsible for causing liver abscesses in cattle, is a lactate utilizer whose abundance can increase up to $10^6$ to $10^7$ cells/g when animals are fed grain diets (Coe et al., 1999; Tan, Nagaraja, & Chengappa, 1994).
4.1 Rationale and Approach

In ruminal microbial communities, bacteria represent the most abundant and diverse group of microorganisms. They play an essential role in converting components of ingested feed stuffs into SCFAs, which animal hosts can use as energy (Schwartz, 1969). Amongst the different strategies used to increase ruminant production, feeding ingredients rich in starch has become a standard practice. However, feeding high amounts of starch leads to increased production and accumulation of not only the major ruminal SCFAs, but also of lactate. Elevated ruminal SCFAs and lactate can result in acidosis, a condition that can harm animals, and negatively impact their productivity. In order to design more effective strategies to mitigate the incidence of acidosis, there is a critical need to improve our understand of amylolytic and lactate-utilizing bacteria in the rumen. Indeed, amylolytic bacteria are critical for digestion of starch, while lactate-utilizing bacteria are essential to prevent the accumulation of lactate, which is a common end product of amylolytic bacteria.

Because of the high microbial diversity that exists in the rumen, the ability to assign specific metabolic functions to particular bacterial species has remained a challenge. Indeed, not only does the vast majority of ruminal microbial species remain unknown, but ruminant diets can include a wide variety of substrates, which promotes increased functional and microbial diversity.

The traditional approach to this problem has been the use of cultivation-based methods, i.e. using controlled selective culture conditions to identify microbial isolates with metabolic activities of interest. However, the majority of ruminal bacterial species have so far proven resistant to isolation by growth on culture medium, which may be due
to factors such as the absence of essential nutrients in current medium formulations or to the necessity of co-culturing with other microorganisms to achieve growth.

One alternative to the selection of isolates is to use an enrichment approach. In this case, samples from the environment of interest are used as the culture medium, which will not only provide a pool of microorganisms with the metabolic activity of interest, but also provide essential nutrients as well as microbial partner species that may be necessary for their growth. Enrichment for bacterial species that express a metabolic activity of interest can be achieved by supplementing a set of replicate cultures with a relevant substrate. Since the majority of ruminal microbial species remain uncharacterized, a DNA-based method, such as 16S rRNA profiling, can be used to identify enriched species by comparing the bacterial composition of supplemented and unsupplemented cultures. Similarly, metagenomics can be used to gain further insight on the metabolic potential of enriched species through gene annotation of partially reconstructed genomes.
4.2 Research Objectives

In this context, the main objectives of the research presented in this dissertation were to

- identify bacteria that can grow in batch cultures consisting of rumen fluid supplemented with starch
- determine the metabolic potential of starch-enriched bacterial species using metagenomics
- identify bacteria that can grow in batch cultures consisting of rumen fluid supplemented with lactate
- determine the metabolic potential of lactate-enriched bacterial species using metagenomics
Figure 1.1. Amylolytic, maltose-fermenting, glucose-fermenting and lactic acid fermenting bacteria involved in the fermentation of high-concentrate diets (Nagaraja & Titgemeyer, 2007).
Table 1.1. PCR and Non-PCR Based Molecular Approaches.

<table>
<thead>
<tr>
<th>PCR-based molecular techniques</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturing/temperature gradient gel electrophoresis</td>
<td>(Muyzer, 1999)</td>
</tr>
<tr>
<td>Single-strand-conformation polymorphism (SSCP)</td>
<td>(Lee, Zo, &amp; Kim, 1996)</td>
</tr>
<tr>
<td>Restriction fragment length polymorphisms (RFLP)</td>
<td>(LAGUERRE, RIGOTTIER-GOIS, &amp; LEMANCEAU, 1994)</td>
</tr>
<tr>
<td>Terminal restriction fragment length polymorphisms (T-RFLP)</td>
<td>(Dunbar, Dragon, Lee, &amp; Baserga, 2000)</td>
</tr>
<tr>
<td>Quantitative PCR (qPCR)</td>
<td>(Takai &amp; Horikoshi, 2000)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-PCR-based molecular techniques</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluorescence in situ hybridization (FISH) and microarray</td>
<td>(Bodrossy &amp; Sessitsch, 2004; Bottari, Ercolini, Gatti, &amp; Neviani, 2006)</td>
</tr>
<tr>
<td>Raman microspectroscopy</td>
<td>(Huang, Li, Jarvis, Goodacre, &amp; Banwart, 2010)</td>
</tr>
<tr>
<td>Nano-scale secondary ion mass spectrometry (NanoSIMS)</td>
<td>(Herrmann et al., 2007)</td>
</tr>
<tr>
<td>Next-generation sequencing (NGS) methods pyrosequencing</td>
<td>(Nowrousian et al., 2010)</td>
</tr>
<tr>
<td>Metagenomics</td>
<td>(XU, 2006)</td>
</tr>
<tr>
<td>Metatranscriptomics</td>
<td>(Poretsky et al., 2009)</td>
</tr>
<tr>
<td>Metaproteomics</td>
<td>(Wilmes &amp; Bond, 2004)</td>
</tr>
</tbody>
</table>
Table 1.2. Overview of amount of starch fermented in the rumen when the livestock are supplemented with different starch sources.

<table>
<thead>
<tr>
<th>Grain</th>
<th>Starch (%)</th>
<th>Rumen digestability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corn (Herrera-Saldana, Huber, &amp; Poore, 1990; Huntington, Harmon, &amp; Richards, 2006)</td>
<td>76.0</td>
<td>72-89.9</td>
</tr>
<tr>
<td>Sorghum (Herrera-Saldana et al., 1990; Huntington et al., 2006)</td>
<td>71.3</td>
<td>60-78.4</td>
</tr>
<tr>
<td>Wheat (Herrera-Saldana et al., 1990; Huntington et al., 2006)</td>
<td>70.3</td>
<td>88.3-88.1</td>
</tr>
<tr>
<td>Barley (Herrera-Saldana et al., 1990; Huntington et al., 2006)</td>
<td>64.3</td>
<td>80.7-84.6</td>
</tr>
<tr>
<td>Oats (Herrera-Saldana et al., 1990; Huntington et al., 2006)</td>
<td>58.1</td>
<td>92.7-94.0</td>
</tr>
<tr>
<td>Yucca (Gómez, Posada, &amp; Olivera, 2016)</td>
<td>80.0</td>
<td>91.0</td>
</tr>
</tbody>
</table>
References


Carberry, C. A., Kenny, D. A., Han, S., McCabe, M. S., & Waters, S. M. (2012). Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial


analytical tool in biogeochemistry and soil ecology: A review article. *Soil Biology and Biochemistry, 39*(8), 1835-1850. doi:https://doi.org/10.1016/j.soilbio.2007.03.011


doi:10.1016/j.anaerobe.2015.12.003

doi:10.1007/s00248-017-0940-z


doi:10.1371/journal.pone.0020521


to Solid Digesta in the Rumen of Cows. *PLOS ONE, 8*(11), e78507.
doi:10.1371/journal.pone.0078507

doi:10.3389/fmicb.2019.00649


clinical microbiology laboratories. Clin Microbiol Infect, 14(10), 908-934. 

XU, J. (2006). INVITED REVIEW: Microbial ecology in the age of genomics and 
metagenomics: concepts, tools, and recent advances. Molecular Ecology, 15(7), 1713-
1731. doi:10.1111/j.1365-294X.2006.02882.x

Youssef, N., Sheik, C. S., Krumholz, L. R., Najar, F. Z., Roe, B. A., & Elshahed, M. S. 
(2009). Comparison of species richness estimates obtained using nearly complete 
fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based 
doi:10.1128/aem.00592-09

<em>rrs</em> Genes for Use in Fingerprinting of Microbial Communities by 
PCR-Denaturing Gradient Gel Electrophoresis. Applied and environmental microbiology, 
70(8), 4800. doi:10.1128/AEM.70.8.4800-4806.2004

doi:10.1016/j.biortech.2012.11.073

equal than others: the role of "keystone" species in the degradation of recalcitrant 

Dietary Forage to Concentrate Ratios on Dynamic Profile Changes and Interactions of
Ruminal Microbiota and Metabolites in Holstein Heifers. *Front Microbiol*, 8, 2206.
CHAPTER 2

IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE STARCH UTILIZING BACTERIA FROM THE RUMEN OF BEEF CATTLE
ABSTRACT

Starch is the major energy component in grain based-diets. Feedlot beef and dairy cattle are fed with grain diets to maximize their performance. Rumen amylolytic or starch metabolizing bacteria predominate when the animals are fed with grain-based diets. Since the specific role of amylolytic bacteria have yet to be clearly elucidated, the primary objective of this research was to identify and determine the metabolic functions of previously uncharacterized amylolytic bacteria. Our general approach consisted of culturing rumen fluid using anaerobic bioreactors for a period of 14 days. Four experimental trials, each performed with a different rumen fluid donor, designated as Trial 1, Trial 2, Trial 3, and Trial 4, were conducted during this study. Four replicate cultures, each consisting of approximately 2.2 L of rumen fluid per laboratory-scale bioreactors were used. Three of the replicate cultures were supplemented with starch (11.4g/L, ADM Corn processing, Clinton, Iowa), while the remaining culture was not supplemented with any substrate (Control). Cultures were maintained for 14 days followed by sample collection on day-7 and day-14. Four species-level OTUs, corresponding to genus Ruminobacter and Prevotella were enriched on days 7 and 14. Prevotella-affiliated OTUs (SD_Bt-00010 and SD_Bt-00966) were then selected for Shotgun metagenomics analysis. This analysis resulted in the identification of enzymes that were predicted to be involved in starch metabolism. Further characterizing these Prevotella-affiliated OTUs may contribute towards a better understanding of starch metabolism.
1. Introduction

As a result of their ability to produce high quality protein products such as milk and meat from plant fibers, which cannot be efficiently digested by humans, ruminants have so far played a vital role throughout mankind history, and they are expected to continue to do so in the foreseeable future (Flachowsky, Meyer, & Sudekum, 2017) (Flint, Scott, Duncan, Louis, & Forano, 2012). In these herbivores, digestion of feed takes place in the rumen, the largest compartment of a four-chambered stomach, through the combined metabolic activities of resident microbial symbionts which precede digestion by host enzymes (Dehority, 2003) (Hungate, 1966). Ruminal microorganisms, consisting of a diverse array of bacteria, methanogens, protozoa, and fungi, are organized into complex communities that work synergistically to ferment ingested feedstuffs, producing short chain fatty acids (SCFAs) and microbial proteins as end products that provide energy and amino acids for their host (S. Wang et al., 2017) (Lengowski et al., 2016).

While ruminants have evolved to utilize structural polysaccharides as their primary source of energy, the inclusion of starch-rich feedstuffs, such as maize or cereal grains, is a common practice in intensive ruminant livestock production systems (Huntington, Harmon, & Richards, 2006) (Huntington, 1997). Indeed, since starch is a more readily available source of energy compared to plant fiber polysaccharides, it provides a more efficient means of meeting the energy requirements of high producing ruminants (Gómez, Posada, & Olivera, 2016). However, faster rates of fermentation for starch compared to plant fiber components can result in rapid accumulation of short chain fatty acids (SCFAs), which can breach the buffering capacity limits of the rumen, and cause ruminal acidosis (Nagaraja & Titgemeyer, 2007) (Goad, Goad, & Nagaraja, 1998).
Indeed, while fluctuations in ruminal pH are normally observed in healthy individuals, acidotic ruminal pH conditions that are maintained for extended periods of time can have a detrimental effect on the health of affected individuals (Snyder & Credille, 2017). As acidosis not only impairs normal functions of the digestive tract, but can also lead to further complications such as bloat, diarrhea, liver abscesses, and laminitis, it can have a negative impact on animal performance and health (Vyas et al., 2014), resulting in economic losses for ruminant livestock procedures (Grohn & Bruss, 1990).

In intensive production systems, acidosis is most commonly managed by gradually increasing the proportion of starch in the diet over the course of an adaptation period (Owens, Secrist, Hill, & Gill, 1998). While implementation of this practice and other management strategies has greatly helped in reducing the impact of acidosis on the industry, it still remains a concern for animal health and profitability (NejashAbdela, 2016). Since one of the main causes of acidosis is the rapid accumulation of ruminal SCFAs from microbial digestion of high starch diets, a great deal of effort has been devoted to the identification of ruminal microorganisms that participate in this process (T. J. DeVries, T. Schwaiger, Beauchemin, & PennerB, 2014). One reported strategy has been to investigate the response or dynamics of ruminal bacterial communities during subacute ruminal acidosis (SARA), a reversible state of pH depression that can further develop into acidosis (Plaizier, Li, Tun, & Khafipour, 2016). In animal nutritional models, induction of SARA using starch-rich feedstuffs was found to be associated with reduced bacterial richness and diversity (Plaizier et al., 2017) (S. Li et al., 2012) (Hook et al., 2011), (Khafipour, Li, Plaizier, & Krause, 2009), with observed reductions in
Bacteroidetes and increases in Firmicutes (Khafipour et al., 2009), as well as changes in a number of well characterized bacterial species that are known to metabolize starch.

Considering that typically only a fraction of microorganisms identified in rumen samples correspond to valid species, it is generally acknowledged that the vast majority of ruminal symbionts remain to be characterized (Creevey, Kelly, Henderson, & Leahy, 2014). Based on this assessment, it was hypothesized that currently unknown ruminal microorganisms include amylolytic bacteria that have yet to be identified. Using a batch-culture system, candidate starch utilizers were enriched from beef cow rumen fluid, then characterized as Operational Taxonomic Units (OTUs) using a 16S rRNA-based approach. Metagenomics analyses were conducted to assess the metabolic potential of candidate ruminal bacterial species, which revealed coding sequences for enzymes that participate in pathways predicted to metabolize starch into SCFAs. Together, the results presented in this report suggest that the bacterial species corresponding to these OTUs function as starch utilizers in the rumen of beef cows.

2. Materials and Methods

2.1 Sample collection and in vitro rumen culture experiments

All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at South Dakota State University. Fresh rumen fluid was obtained by manually squeezing ruminal digesta collected from fistulated beef cows maintained at the South Dakota State University Cow-Calf Research Facilities. The diet of these animals consisted of pasture hay or haylage during the period when rumen fluid was
collected (March 2016 to November 2016). Four experimental trials, each performed with a different rumen fluid donor, designated as Trial 1, Trial 2, Trial 3, and Trial 4, were conducted during this study, with the following procedure followed for each trial. Four replicate cultures, each consisting of approximately 2.2 L of rumen fluid per laboratory-scale bioreactors (Chemglass), were set up within one hour of collection. Three of the replicate cultures were supplemented with starch (11.4g/L, ADM Corn processing, Clinton, Iowa, USA), while the remaining culture was not supplemented with any substrate (Control). Treatment and control cultures were maintained under anaerobic conditions at a constant physiological temperature (38°C), with continuous agitation at 150 rpm using a Rushton-style impeller fixed to a stirring shaft built into the bioreactor. Each bioreactor allowed for approximately 0.8 L of headspace, and was equipped with a flexible plastic tube for exhaust of excess biogas to prevent pressure build-up. Culture samples (approximately 15 mL/sample) were collected on days 7 (D7) and 14 (D14) to assess their respective bacterial composition. Samples from the rumen inoculum (D0) and cultures were stored frozen at -20°C until analyzed.

2.2 Microbial genomic DNA purification and PCR amplification of the 16S rRNA gene

Genomic DNA was extracted from each sample by a repeated bead beating plus column method as previously described (Yu & Morrison, 2004). Briefly, 250µL of rumen sample was lysed in extraction buffer (0.5 M NaCl, 50 mM Tris.HCl, 50Mm EDTA, 4% SDS) by bead beating followed by sequential extraction with 10M ammonium acetate then isopropanol precipitation. Recovered nucleic acids were then purified using the QIAamp DNA Stool Kit (QIAGEN, Hilden, Germany) following the manufacturer’s recommended protocol. PCR was performed using the Phusion Taq DNA polymerase
(ThermoFischer Scientific, Waltham, MA, USA) on a 2720 Thermo Cycler (ThermoFischer Scientific), with the 27F (Edwards, Rogall, Blocker, Emde, & Bottger, 1989) and 519R (Lane et al., 1985) primers to target the V1-V3 regions of the 16S rRNA gene. PCR reactions consisted of a ‘hot start’ (98°C, 3 min), followed by 35 consecutive cycles of denaturation (98°C, 30 sec), annealing (50°C, 30 sec), and elongation (72°C, 30 sec), then by a final elongation period (72°C, 10 min). Quality of PCR amplicons (expected approximate length of 500 bp) was determined by agarose gel electrophoresis, with recovery of PCR-generated DNA using the QiaexII Gel extraction kit (Qiagen, Hilden, Germany). Gel purified amplicons from each sample were then submitted as template for High Throughput Sequencing using an Illumina Miseq(2×300) platform (Molecular Research DNA, Shallowater, TX).

2.3 Bioinformatic analysis for 16S rRNA gene-based composition analysis

Unless specified, datasets were analyzed using custom written Perl scripts (available upon request). Raw bacterial 16S rRNA gene V1-V3 amplicon sequences were provided by Molecular Research DNA as contigs assembled from overlapping MiSeq 2x300 paired-end reads from the same flow cell clusters. Reads were then selected to meet the following criteria: presence of both intact 27F (forward) and 519R (reverse) primer nucleotide sequences, length between 400 and 580 nt, and a minimal quality threshold of no more than 1% of nucleotides with a Phred quality score lower than 15.

Following quality screens, sequence reads were aligned, then clustered into Operational Taxonomic Units (OTUs) at a genetic distance cutoff of 5% sequence dissimilarity (St-Pierre & Wright, 2015). While 3% is the most commonly used clustering cutoff for 16S rRNA, it was originally recommended for full length sequences, and may
not be suitable for the analysis of specific sub-regions since nucleotide sequence variability is not constant across the entire length of the 16S rRNA gene. In this context, if 3% is a commonly accepted clustering cutoff for V4 or V4-V5 regions, which are the least variable of the hypervariable regions, then a higher cutoff should be used for the V1-V3 region, since V1 is the most variable region of the 16S rRNA gene. OTUs were screened for DNA sequence artifacts using the following methods. Chimeric sequences were first identified with the chimera.uchime and chimera.slayer commands from the MOTHUR open source software package (Schloss et al., 2009). Secondly, the integrity of the 5’ and 3’ ends of OTUs was evaluated using a database alignment search-based approach; when compared to their closest match of equal or longer sequence length from the NCBI nt database, as determined by BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1997), OTUs with more than five nucleotides missing from the 5’ or 3’ end of their respective alignments were discarded as artifacts. Single read OTUs were subjected to an additional screen, where only sequences that had a perfect or near perfect match to a sequence in the NCBI nt (National Center for Biotechnology Information) database were kept for analysis, i.e. that the alignment had to span the entire sequence of the OTU, and a maximum of 1% of dissimilar nucleotides was tolerated. The number of non-chimeric and artifact-free sequence reads used for analysis in each trial is shown in Table 2.1.

After removal of sequence chimeras and artifacts, RDP Classifier (Ribosomal Database Project) (Q. Wang, Garrity, Tiedje, & Cole, 2007) and BLAST (Altschul et al., 1997) were used for taxonomic assignment of valid OTUs. The List of Prokaryotic
Names with Standing in Nomenclature (LPSN) was also consulted for information on valid species belonging to taxa of interest (Parte, 2014).

2.4 Metagenomic analysis

A select number of samples exhibiting high levels of enrichment for candidate starch utilizing OTUs were further analyzed using a metagenomics approach. Purified microbial genomic DNA (extracted as described above) was used directly as template for high throughput sequencing using an Illumina Miseq(2×250) platform (Molecular Research DNA, Shallowater, TX). Raw sequence reads of 200 bp in length or longer were selected using custom written Perl scripts for building of genomic contigs using the *de novo* assembly program ABySS (Simpson et al., 2009). ABySS was run on a Linux-based high-performance computing cluster maintained by the University Networking and Research Computing Group at South Dakota State University, with k-mer length set at 64 and all other parameters used at their respective default values.

ABySS-assembled contigs of 4000 bp or greater were selected for further analysis by gene annotation. Coding sequences (CDS) were identified and annotated using RAST (Rapid Annotations using Subsystems Technology) (Aziz et al., 2008), with assignment of CDS to metabolic pathways performed using KEGG pathways as a model reference (Kanehisa & Goto, 2000). Further analysis of select CDS was performed against the NCBI nr database using BLASTp (Altschul et al., 1997) to identify their respective closest homologs.
3. Results

3.1 Comparative analysis of bacterial communities from donor rumen fluid

Analysis of the rumen fluid inocula before culturing with starch revealed that bacteria affiliated to the phyla Bacteroidetes and Firmicutes were overall the most abundant, representing 24.8 - 68.1% and 24.5 - 68.2% of sequence reads, respectively, with their combined totals ranging between 73.5 and 93.0% (Figure 2.1). Prevotellaceae was the predominant Bacteroidetes family (14.4 - 53.1%), while Lachnospiraceae and Ruminococcaceae were the most highly represented Firmicutes families (1.8 - 16.8% and 1.6 - 22.2%, respectively).

A comparative analysis further revealed that 54 Operational Taxonomic Units (OTUs) were shared amongst all donors, representing 6.9 - 24.0% of sequence reads in individual donor rumen fluid samples, while the abundance of sequences that were specific or unique to individual donors ranged between 20.9% and 46.1% per sample (Table 2.2). Amongst the 54 shared OTUs, 10 OTUs were identified as the most abundant, with their respective individual abundance ranging between 0.06% and 5.29% across all samples (Table 2.3). Only one OTU showed species-level sequence identity to its closest valid taxon (Bt-00074; 98.3%), while the other abundant OTUs were found to be more distant (84.6 – 92.4%), and likely corresponded to novel bacterial species. Seven of the most abundant OTUS were affiliated to Bacteroidetes, of which six were predicted to be members of the genus Prevotella, while the remaining three OTUs belonged to Firmicutes.
3.2 Identification of candidate bacterial starch utilizers from rumen fluid

When compared to their respective culture controls without substrate, three distinct OTUs were found to be the most enriched bacteria in starch-supplemented rumen fluid cultures (Table 2.4). From the Trial 1 donor, OTU SD_Bt-01020 was in highest abundance in starch supplemented cultures at day 7, at an average value that was 1206X greater than the non-supplemented control. The closest valid relative to SD_Bt-01020 was *Ruminobacter amylophilus*, a well characterized starch digester (Anderson, 1995), with a sequence identity typically considered to be within species-level range for 16S rRNA gene-based analyses. OTU SD_Bt-00966 was identified as a candidate starch utilizer from Trial 2. Compared to other enriched OTUs in this study, it exhibited the most consistent level of enrichment across replicates and time points. Intriguingly, the abundance of this OTU was already elevated in the rumen inoculum of Trial 2, prior to culturing with starch. *Prevotella copri*, originally isolated from human feces (Hayashi, Shibata, Sakamoto, Tomita, & Benno, 2007), was identified as its closest bacterial species, but at a degree of 16S rRNA gene sequence identity more consistent with genus-level rather than species-level cutoffs. Lastly, the remaining candidate starch utilizing OTU identified was SD_Bt-00010. It was enriched from two separate donors (Trial 3 and Trial 4), and showed the highest abundance averages amongst sets of starch supplemented replicate cultures in this study. SD_Bt-00010 was found to be within species-level sequence identity to *Prevotella albensis*, for which $\alpha$-amylase activity has previously been reported (Avgustin, Wallace, & Flint, 1997).
3.3 Exploring the metabolic potential of OTUs identified as candidate starch utilizers

We postulated that sequence datasets generated from samples with high abundance of an OTU would have higher representation of genomic sequences from this OTU. Consequently, higher sequence representation or coverage would then be favorable for the assembly of contigs that would be of sufficient length to allow for prediction of full-length coding sequences and improved accuracy of gene annotation. From the three OTUs that were enriched in the presence of starch, we focused our attention on the *Prevotella*-affiliated OTUs (SD_Bt-00010 and SD_Bt-00966), since *Prevotella*-affiliated symbionts are typically the most abundant and diverse group of uncharacterized bacteria in the rumen of cattle.

The most highly enriched samples for SD_Bt-00010 (Trial 3: 67.9% and Trial 4: 74.0%) and for SD_Bt-00966 (Trial 2: 24.9%), were then selected for shotgun sequencing, which generated a total of 9.5 X 10^6, 9.16 X 10^6 and 8.4 X 10^6 paired end reads, respectively. After de novo assembly, contigs were filtered for a minimal length of 4,000 nt, which resulted in the selection of 436 (Trial 3), 397 (Trial 4), and 983 (Trial 2) contigs for further analysis. The maximum contig length from each sample was found to be 68,555 nt (Trial 3), 181,113 nt (Trial 4), and 28,496 nt (Trial 2). The RAST online tool was then used for predicting amino acid coding regions (CDS) from the selected contigs, as well as for performing gene annotation. For the Trial 2 contig set, 974 of 7,422 CDS were annotated successfully, compared to 916 of 5,407 CDS and 1,008 of 4,848 CDS for Trials 3 and 4, respectively.

Predicted metabolic functional categories from gene annotation included carbohydrates, proteins, amino acids and derivatives, DNA and RNA, phosphorus,
potassium and sulfur, regulation and cell signaling, as well as stress response (Figure 2.2). Based on the conditions of the culturing assay, our primary interest was to identify genes involved in carbohydrate metabolism; 345 CDS were assigned to this function from the Trial 2 contig set, compared to 377 CDS for the Trials 3 and 4 datasets.

To gain further insight, data mining for specific functions in carbohydrate metabolism was performed, resulting in the identification of enzymes predicted to be associated with metabolic activities of amylolytic bacteria (Figures 2.3, 2.4, and 2.5). Coding sequences for α-amylase (EC 3.2.1.1), which catalyzes the hydrolysis of α-1,4-glycosidic bonds in starch to produce maltose, as well as for α-glucosidase (EC 3.2.1.20), which catalyzes the conversion of maltose into glucose, were identified for SD_Bt-00010 and for SD_Bt-00966. All enzymes of the glycolysis pathway were also found for both OTUs, indicating that glucose released from the hydrolysis of starch would be predicted to be metabolized to pyruvate and yield ATP in these candidate bacterial species (Figures 2.3, 2.4, and 2.5).

For both OTUs, enzymes involved in further metabolizing pyruvate into SCFAs were also identified (Figures 2.3, 2.4, and 2.5). Two outcomes were predicted for pyruvate, with either a conversion to lactate through lactate dehydrogenases (EC 1.1.1.27 and EC 1.1.1.28) or the production of formate and acetyl-CoA through the activity of pyruvate formate-lyase (EC 2.3.1.54). Formate could then be oxidized into CO₂ by formate dehydrogenase, while acetyl-CoA could be metabolized sequentially into acetyl phosphate (phosphate acetyltransferase, EC 2.3.1.8), acetate (acetate kinase, EC 2.7.2.1), aldehyde (aldehyde dehydrogenase, EC 1.2.1.3), then ethanol (alcohol dehydrogenases, EC 1.1.1.1).
In addition to glucose fermentation pathways, enzymes involved in glycogen synthesis and utilization were also identified in SD_Bt-00010 and SD_Bt-00966 affiliated contigs, indicating that glucose released from starch hydrolysis could also be stored as glycogen. Enzymes predicted to be involved in this function included glycogen synthase (EC 2.4.1.21), which converts ADP-glucose to amylose, as well as 1,4-alpha-glucan branching enzyme (EC 2.4.1.18), which generates glycogen from amylose. Glucose stored as glycogens would then be predicted to be mobilized through the sequential actions of glycogen phosphorylase (EC 2.4.1.1) and phosphoglucomutase (EC 5.4.2.2), which would generate glucose-1-phosphate and glucose-6-phosphate, respectively. The latter could then enter glycolysis to be metabolized into pyruvate and SCFAs.

For amino-related metabolism, contig sets from SD_Bt-00010 and SD_Bt-00966-enriched samples included a high number of coding sequences for proteases and peptidases, as well as transporters for amino acids and oligopeptides (Table 2.8), ranging from 16 to 36 and 30 to 49, respectively. Based on their positioning and close proximity on contigs, coding sequences for transporters suggested they could potentially have represented subunits for multi-protein complexes that function in amino acid transport. Together, these results suggested that candidate bacterial species SD_Bt-00010 and SD_Bt-00966 could acquire amino acids through hydrolysis of proteins present in their immediate environment, followed by transport of amino acids for use in protein synthesis or as substrates for fermentation.
4. Discussion

Carbohydrates in ruminant diets are fermented by the collective actions of a wide array of ruminal microorganisms. In animals raised in intensive management systems, diets typically include high levels of grain to favor increased performance yields (Huntington et al., 2006) (Owens, Zinn, & Kim, 1986). Typically, different groups of bacteria, namely cellulolytic and amylolytic, each specialize in the digestion of either cellulose or starch, respectively (Nafikov & Beitz, 2007). Since starch is one of the most abundant carbohydrates in grain, high grain diets then tend to favor amylolytic bacteria (Huntington, 1997) (Mills, France, & Dijkstra, 1999). However, as primary starch utilizers, amylolytic bacteria can be instigators of SARA, since the faster rate of starch breakdown results in greater levels of SCFA production, with a more severe impact on ruminal pH. Thus, further investigation of the metabolic activities of indigenous ruminal amylolytic bacteria is critical to the development and improvement of effective strategies to prevent or minimize the onset and effects of SARA / acidosis (Cotta, 1988) (Morgante, Stelletta, Berzaghi, Gianesella, & Andrighetto, 2007) (T. J. DeVries et al., 2014). However, one critical gap in our knowledge is that most amylolytic ruminal bacteria likely remain to be identified and characterized, as an estimated 95% of ruminal bacterial species have yet to be assigned a function (Creevey et al., 2014).

In this context, the primary objective of the research presented in this report was to identify uncharacterized rumen bacteria that can metabolize starch, then assess their metabolic potential using a metagenomic approach. The bacterial composition of the four rumen samples prior to culturing in the presence of starch (D0), in which Bacteroidetes and Firmicutes were the most abundant phyla and Prevotellaceae were overall the most
predominant family, was consistent with previously reported studies (Pitta et al., 2014) (Lima et al., 2015) (Opdahl, Gonda, & St-Pierre, 2018) (Jami & Mizrahi, 2012) (Jami, White, & Mizrahi, 2014) (Jewell, McCormick, Odt, Weimer, & Suen, 2015). From these complex communities, three main OTUs were found to be in higher relative abundance in ruminal fluid cultures when starch was provided as the only supplemented substrate.

Enrichment for OTU SD_Bt-01020, predicted to be a strain of *Ruminobacter amylophilus* based on 16S rRNA gene sequence comparisons, was maintained to similar levels in both day-7 and day-14 starch-supplemented cultures. *Ruminobacter amylophilus* has been reported as a key amylolytic species involved in starch hydrolysis based on its amylase, amylopectinase, and pullulanase activities (Anderson, 1995), and it has been found to be an excellent model for investigating starch hydrolysis (Petri, Forster, Yang, McKinnon, & McAllister, 2012) (Singh et al., 2014) (Anderson, 1995).

The other two enriched OTUs from this study, SD_Bt-00966 and SD_Bt-00010, were assigned to the lineage *Prevotella*, which has been found to be prevalent in a number of gastro-intestinal environments, including bovine rumen and human gut (Qian et al., 2017) (Jung et al., 2018) (Kovatcheva-Datchary et al., 2015) (Bhute et al., 2016) (Girard, Tromas, Amyot, & Shapiro, 2017). While SD_Bt-00966 and SD_Bt-00010 were both identified as starch utilizers, they may belong to different groups within this functional category, as they were each related to different species, *Prevotella copri* and *Prevotella albensis*, respectively. Previously, *Prevotella copri* was reported to be more prominent in humans that ate barley-kernel bread compared to humans eating white wheat flour bread (Kovatcheva-Datchary et al., 2015), while *Prevotella albensis* was found to utilize starch when strains were supplemented with oat spelt xylan,
carboxymethyl cellulose and starch (Avgustin et al., 1997). While 16S rRNA phylogenetic studies can provide a comprehensive picture of bacterial community composition, they are limited in the functional insights that they can provide (Creevey et al., 2014) (Deusch, Tiloca, Camarinha-Silva, & Seifert, 2015). For this reason, a metagenomics analysis was performed on representative high enrichment samples to get an assessment of the metabolic potential of the *Prevotella*-affiliated OTUs SD_Bt-00966 and SD_Bt-00010.

Metagenomics and metatranscriptomics studies have shown that starch metabolism is one of the prominent core pathways in animals fed grain-based diets, with α-amylases, the enzymes that catalyze hydrolysis of α-1,4-glucosidic bonds that link chains of glucose monomers in starch (Janecek, Svensson, & MacGregor, 2014) (Janecek, 1994) (Perry et al., 2007), being encoded by highly expressed genes (Mann, Wetzels, Wagner, Zebeli, & Schmitz-Esser, 2018) (F. Li & Guan, 2017) (L. Wang, Hatem, Catalyurek, Morrison, & Yu, 2013). Accordingly, genes encoding α-amylases were identified in contigs from both SD_Bt-00966 and SD_Bt-00010. Besides members of the *Prevotella* genus, other gut bacterial species can express α-amylase as well as other glycoside hydrolases (GH) that are involved in starch metabolism, including *Bacteroides thetaiotaomicron*, a gram-negative obligate anaerobe that predominates in the human colon (D’Elia & Salyers, 1996) (Arnal, Cockburn, Brumer, & Koropatkin, 2018) (Cho, Cho, Wang, & Salyers, 2001) (Baroroh et al., 2017), as well as several strains of *Bifidobacterium* (phylum *Actinobacteria*) (Duranti et al., 2014) (Ryan, Fitzgerald, & van Sinderen, 2006).
Once glucose has been made available from starch hydrolysis, it can be transported to the intracellular environment of ruminal microorganisms where it can be metabolized or stored as glycogen. Accordingly, CDS for all enzymes of the glycolysis pathway were identified in all contig sets analyzed, indicating that both OTUs had the potential to metabolize glucose into pyruvate (Rodríguez, Sosa, & Rodríguez, 2007). The types of SCFAs that are produced from pyruvate tend to vary depending on the microbial species involved. CDS from the SD_Bt-00010 contig set indicated that this bacterial OTU would be capable of producing D- and L-lactate, as well as ethanol. This is consistent with work from a number of different groups that had identified starch utilizers as lactate producers. High levels of lactate in the rumen, particularly D-lactate because it is not metabolized by the host, can lead to SARA and acidosis, a syndrome that can cause severe health problems in animals, ultimately resulting in economic losses to producers (Counotte & Prins, 1981). These results then suggest that SD_Bt-00010 metabolic activity could contribute to the development of acidosis. Currently, one of the main strategies developed to counter SCFA-mediated low pH has been to increase the activity of certain ruminal bacterial species that can metabolize lactate, such as *Selenomas ruminatum* and *Megasphaera elsdenii*, thus countering its accumulation and preventing the development of acidotic conditions in the rumen (Fecskeova, Piknova, Javorsky, & Pristas, 2010). Based on gene annotation analysis of SD_Bt-00010 contigs, an alternative strategy could be to promote production of ethanol instead of lactate. While not as well studied as the latter, ethanol has been reported as a readily detectable ruminal SCFA in concentrate fed animals (Raun & Kristensen, 2011) (Kristensen, Sehested, Jensen, & Vestergaard, 2007) (K.Pradhan & R.W.Hemken, 1970).
In addition to carbohydrate metabolism, the acquisition of amino acids is a function that is essential for the proliferation of ruminal microorganisms. Based on the predicted enzymatic capabilities from the contig sets analyzed in this study, both enriched OTUs appeared to rely on proteins from their immediate environment as an amino acid source, since proteases, peptidases, and oligopeptide transporters were well represented in annotated coding sequences, while only a limited set of enzymes for de novo synthesis of amino acids were predicted. Thus, in contrast to most ruminal bacterial models, SD_Bt-00010 and SD_Bt-00966 may not rely heavily on the use of ammonia for synthesis of amino acids (Mann et al., 2018) (Hart, Creevey, Hitch, & Kingston-Smith, 2018) (Rodríguez et al., 2007).

5. Conclusion

The batch-culture approach used in this study allowed for the identification of previously uncharacterized starch utilizers from the rumen of beef cows. Based on 16S rRNA gene sequence analysis, three ruminal bacterial OTUs enriched with starch as the only provided substrate were found to be affiliated to the genera Ruminococcus and Prevotella. Further investigations of the Prevotella related OTUs using metagenomics revealed genome-derived coding sequences that could be assembled into predicted pathways that would allow the fermentation of starch into SCFAs. Furthermore, these different OTUs shared other metabolic functions, such as the ability to store glucose as glycogen, as well as to utilize extracellular proteins as a source of amino acids. Together, these characteristics could help define members of a particular subgroup of ruminal
starch utilizers, ultimately contributing to a better understanding of metabolic functions and events that impact the onset of SARA and the incidence of acidosis.
Figure 2.1. Taxonomic composition of rumen samples used as inocula, prior to setting up batch culture enrichment experiments with starch. Each experimental trial (T1, T2, T3, T4) is represented on the horizontal axis, with the relative abundance (%) of each major taxonomic group shown on the vertical axis. Families belonging to the same phylum are represented by different shades of the same color: Firmicutes (green), Bacteroidetes (blue), Proteobacteria (red), Planctomycetes (orange), and other families (violet).
Figure 2.2. RAST gene annotation analysis of contig sets generated from enrichment Trials 2 (blue), 3 (orange) and 4 (grey). The relative abundance (%) of coding sequences for each subsystem level category is shown.
Figure 2.3. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#2 (OTU SD_Bt-00966, affiliated to Prevotella copri). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization. Each enzymes is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 2.5 for corresponding enzyme names.
Figure 2.4. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU SD_Bt-00010, affiliated to Prevotella albensis). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization. Each enzymes is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 2.6 for corresponding enzyme names.
Figure 2.5. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#4 (OTU SD_Bt-00010, affiliated to *Prevotella albensis*). A) starch metabolism and glycolysis pathways; B) glycogen synthesis and mobilization. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 2.7 for corresponding enzyme names.
Table 2.1. Number of high quality-chimera free sequence reads used for OTU analysis in each trial

<table>
<thead>
<tr>
<th>Trial</th>
<th>Number of sequence reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>237,750</td>
</tr>
<tr>
<td>2</td>
<td>243,235</td>
</tr>
<tr>
<td>3</td>
<td>243,316</td>
</tr>
<tr>
<td>4</td>
<td>201,202</td>
</tr>
</tbody>
</table>
Table 2.2: Shared ruminal bacterial OTUs amongst rumen fluid donors used in this study. Abundance is presented as relative abundance (%) of total non-chimeric reads per sample.

<table>
<thead>
<tr>
<th>Trials</th>
<th>Shared OTUs</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1,T2,T3,T4</td>
<td>54</td>
<td>21.8</td>
<td>6.9</td>
<td>21.9</td>
<td>24.0</td>
</tr>
<tr>
<td>T1,T2,T3</td>
<td>247</td>
<td>26.2</td>
<td>8.2</td>
<td>23.1</td>
<td>0</td>
</tr>
<tr>
<td>T1,T2,T4</td>
<td>5</td>
<td>1.0</td>
<td>0.2</td>
<td>0</td>
<td>1.2</td>
</tr>
<tr>
<td>T1,T3,T4</td>
<td>49</td>
<td>2.9</td>
<td>0</td>
<td>4.1</td>
<td>15.4</td>
</tr>
<tr>
<td>T2,T3,T4</td>
<td>12</td>
<td>0</td>
<td>2.6</td>
<td>1.1</td>
<td>4.5</td>
</tr>
<tr>
<td>T1,T2</td>
<td>74</td>
<td>3.1</td>
<td>1.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T1,T3</td>
<td>555</td>
<td>18.8</td>
<td>0</td>
<td>21.3</td>
<td>0</td>
</tr>
<tr>
<td>T1,T4</td>
<td>21</td>
<td>0.7</td>
<td>0</td>
<td>0</td>
<td>4.3</td>
</tr>
<tr>
<td>T2,T3</td>
<td>202</td>
<td>0</td>
<td>34.5</td>
<td>5.5</td>
<td>0</td>
</tr>
<tr>
<td>T2,T4</td>
<td>5</td>
<td>0</td>
<td>0.08</td>
<td>0</td>
<td>2.1</td>
</tr>
<tr>
<td>T3,T4</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>2.1</td>
<td>9.6</td>
</tr>
<tr>
<td>unique</td>
<td>-</td>
<td>25.5</td>
<td>46.1</td>
<td>20.9</td>
<td>38.2</td>
</tr>
<tr>
<td>Observed OTUs</td>
<td>-</td>
<td>2569</td>
<td>1651</td>
<td>3547</td>
<td>375</td>
</tr>
</tbody>
</table>
Table 2.3. Most abundant# ruminal bacterial OTUs that were shared amongst all rumen fluid donors used in this study. Abundance is presented as relative abundance (%) of total non-chimeric reads per sample.

<table>
<thead>
<tr>
<th>OTU</th>
<th>Trial 1</th>
<th>Trial 2</th>
<th>Trial 3</th>
<th>Trial 4</th>
<th>Closest relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bt-00026F</td>
<td>1.99</td>
<td>0.12</td>
<td>3.13</td>
<td>3.51</td>
<td><em>Christensenella massiliensis</em> (84.6%)</td>
</tr>
<tr>
<td>Bt-00028F</td>
<td>0.10</td>
<td>0.21</td>
<td>1.71</td>
<td>1.36</td>
<td><em>Sporobacter termitidis</em> (87.2%)</td>
</tr>
<tr>
<td>Bt-00030B</td>
<td>1.66</td>
<td>0.92</td>
<td>1.34</td>
<td>0.97</td>
<td><em>Prevotella ruminicola</em> (92.4%)</td>
</tr>
<tr>
<td>Bt-00035F</td>
<td>1.01</td>
<td>0.09</td>
<td>1.28</td>
<td>0.78</td>
<td><em>Neglecta timonensis</em> (85.3%)</td>
</tr>
<tr>
<td>Bt-00047B</td>
<td>2.55</td>
<td>1.21</td>
<td>0.30</td>
<td>0.19</td>
<td><em>Prevotella ruminicola</em> (91.1%)</td>
</tr>
<tr>
<td>Bt-00051B</td>
<td>0.84</td>
<td>0.10</td>
<td>2.15</td>
<td>0.97</td>
<td><em>Odoribacter splanchnicus</em> (82.5%)</td>
</tr>
<tr>
<td>Bt-00071B</td>
<td>0.80</td>
<td>1.34</td>
<td>0.25</td>
<td>0.39</td>
<td><em>Prevotella ruminicola</em> (98.3%)</td>
</tr>
<tr>
<td>Bt-00074B</td>
<td>5.29</td>
<td>0.51</td>
<td>0.51</td>
<td>0.97</td>
<td><em>Prevotella ruminicola</em> (91.4%)</td>
</tr>
<tr>
<td>Bt-00118B</td>
<td>0.46</td>
<td>0.37</td>
<td>1.25</td>
<td>0.19</td>
<td><em>Prevotella ruminicola</em> (90.8%)</td>
</tr>
<tr>
<td>Bt-00119B</td>
<td>0.12</td>
<td>0.06</td>
<td>2.40</td>
<td>0.98</td>
<td><em>Prevotella brevis</em> (90.8)</td>
</tr>
<tr>
<td>Total</td>
<td>14.82</td>
<td>4.93</td>
<td>14.32</td>
<td>10.33</td>
<td></td>
</tr>
</tbody>
</table>

# OTUs found at a relative abundance of at least 1.0% in at least one of the rumen fluid donors.

F. Firmicutes-affiliated OTU.

B: Bacteroidetes-affiliated OTU.
Table 2.4. Enrichment of OTUs in response to starch. Abundance is presented as relative abundance (%) of total non-chimeric reads per sample. D0 (prior to starch enrichment), C-D7 and C-D14 (Control culture day-7 and day-14), S-D7 and S-D14 (Starch culture day-7 and day-14).

<table>
<thead>
<tr>
<th>Donor</th>
<th>OTU</th>
<th>D0</th>
<th>C-D7</th>
<th>S-D7*</th>
<th>C-D14</th>
<th>S-D14*</th>
<th>Closest relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>Bt-01020</td>
<td>0.05</td>
<td>0.03</td>
<td>36.19 ± 5.81</td>
<td>0.15</td>
<td>21.71 ± 11.3</td>
<td>R. amylophilus (97%)</td>
</tr>
<tr>
<td>Trial 2</td>
<td>Bt-00966</td>
<td>19.9</td>
<td>0.25</td>
<td>24.85 ± 3.64</td>
<td>0.35</td>
<td>27.69 ± 4.11</td>
<td>P. copri (91%)</td>
</tr>
<tr>
<td>Trial 3</td>
<td>Bt-00010</td>
<td>0.22</td>
<td>0.03</td>
<td>42.99 ± 21.44</td>
<td>1.13</td>
<td>5.42 ± 2.98</td>
<td>P. albensis (98%)</td>
</tr>
<tr>
<td>Trial 4</td>
<td>Bt-00010</td>
<td>0.11</td>
<td>0.25</td>
<td>33.26 ± 16.83</td>
<td>0.24</td>
<td>70.71 ± 2.44</td>
<td>P. albensis (98%)</td>
</tr>
</tbody>
</table>
Table 2.5. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU SD_Bt-00966, affiliated to *Prevotella albensis*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa#</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-amylase (EC 3.2.1.1)</td>
<td>631</td>
</tr>
<tr>
<td>α-glucosidase (EC 3.2.1.20)</td>
<td>714</td>
</tr>
<tr>
<td>Glucokinase (EC 2.7.1.2)</td>
<td>156</td>
</tr>
<tr>
<td>Glucose-6-phosphate isomerase (EC 5.3.1.9)</td>
<td>453</td>
</tr>
<tr>
<td>Phosphofructo kinase (EC 2.7.1.11)</td>
<td>94</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase class II (EC 4.1.2.13)</td>
<td>468</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)</td>
<td>369</td>
</tr>
<tr>
<td>Phosphoglycerate kinase (EC 2.7.2.3)</td>
<td>33</td>
</tr>
<tr>
<td>Phosphoglycerate mutase (EC 5.4.2.12)</td>
<td>199</td>
</tr>
<tr>
<td>Phosphogluco mutase (EC 5.4.2.2)</td>
<td>314</td>
</tr>
<tr>
<td>Enolase (EC 4.2.1.11)</td>
<td>440</td>
</tr>
<tr>
<td>Pyruvate kinase (EC 2.7.1.40)</td>
<td>480</td>
</tr>
<tr>
<td>Pyruvate formate-lyase (EC 2.3.1.54)</td>
<td>258</td>
</tr>
<tr>
<td>Phosphate acetyltransferase (EC 2.3.1.8)</td>
<td>507</td>
</tr>
<tr>
<td>Acetate kinase (EC 2.7.2.1)</td>
<td>411</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase (EC 1.2.1.3)</td>
<td>1058</td>
</tr>
<tr>
<td>Alcohol dehydrogenase (EC 1.1.1.1)</td>
<td>348</td>
</tr>
<tr>
<td>Glycogen phosphorylase (EC 2.4.1.1)</td>
<td>851</td>
</tr>
<tr>
<td>Glycogen synthase (EC 2.4.1.21)</td>
<td>70</td>
</tr>
<tr>
<td>L-lactate dehydrogenase (EC 1.1.1.27)</td>
<td>314</td>
</tr>
<tr>
<td>Formate dehydrogenase (EC 1.17.1.9)</td>
<td>136</td>
</tr>
<tr>
<td>1,4-α-glucan branching enzyme (EC 2.4.1.18)</td>
<td>750</td>
</tr>
</tbody>
</table>

*# aa: length in amino acids*
Table 2.6. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU SD_Bt-00010, affiliated to *Prevotella albensis*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa#</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-amylase (EC 3.2.1.1)</td>
<td>559</td>
</tr>
<tr>
<td>α-glucosidase (EC 3.2.1.20)</td>
<td>538</td>
</tr>
<tr>
<td>Glucokinase (EC 2.7.1.2)</td>
<td>244</td>
</tr>
<tr>
<td>Glucose-6-phosphate isomerase (EC 5.3.1.9)</td>
<td>440</td>
</tr>
<tr>
<td>Phosphofructo kinase (EC 2.7.1.11)</td>
<td>553</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase class II (EC 4.1.2.13)</td>
<td>186</td>
</tr>
<tr>
<td>Glycerldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)</td>
<td>339</td>
</tr>
<tr>
<td>Phosphoglycerate kinase (EC 2.7.2.3)</td>
<td>393</td>
</tr>
<tr>
<td>Phosphoglycerate mutase (EC 5.4.2.12)</td>
<td>552</td>
</tr>
<tr>
<td>Enolase (EC 4.2.1.11)</td>
<td>471</td>
</tr>
<tr>
<td>Pyruvate kinase (EC 2.7.1.40)</td>
<td>485</td>
</tr>
<tr>
<td>Pyruvate formate-lyase (EC 2.3.1.54)</td>
<td>682</td>
</tr>
<tr>
<td>Phosphate acetyltransferase (EC 2.3.1.8)</td>
<td>330</td>
</tr>
<tr>
<td>Acetate kinase (EC 2.7.2.1)</td>
<td>443</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase (EC 1.2.1.3)</td>
<td>494</td>
</tr>
<tr>
<td>Alcohol dehydrogenase (EC 1.1.1.1)</td>
<td>405</td>
</tr>
<tr>
<td>Glycogen phosphorylase (EC 2.4.1.1)</td>
<td>752</td>
</tr>
<tr>
<td>Glycogen synthase (EC 2.4.1.21)</td>
<td>482</td>
</tr>
<tr>
<td>L-lactate dehydrogenase (EC 1.1.1.27)</td>
<td>319</td>
</tr>
<tr>
<td>D-lactate dehydrogenase (EC 1.1.1.28)</td>
<td>324</td>
</tr>
<tr>
<td>Formate dehydrogenase (EC 1.17.1.9)</td>
<td>931</td>
</tr>
<tr>
<td>1,4-α-glucan branching enzyme (EC 2.4.1.18)</td>
<td>1983</td>
</tr>
<tr>
<td>Phosphogluco mutase (EC 5.4.2.2)</td>
<td>564</td>
</tr>
</tbody>
</table>

# aa: length in amino acids
Table 2.7. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#4 (OTU SD_Bt-00010, affiliated to *Prevotella albensis*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa#</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-amylase (EC 3.2.1.1)</td>
<td>429</td>
</tr>
<tr>
<td>α-glucosidase (EC 3.2.1.20)</td>
<td>538</td>
</tr>
<tr>
<td>Glucokinase (EC 2.7.1.2)</td>
<td>304</td>
</tr>
<tr>
<td>Glucose-6-phosphate isomerase (EC 5.3.1.9)</td>
<td>440</td>
</tr>
<tr>
<td>Phosphofructo kinase (EC 2.7.1.11)</td>
<td>553</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase class II (EC 4.1.2.13)</td>
<td>300</td>
</tr>
<tr>
<td>Glycerladehyde-3-phosphate dehydrogenase (EC 1.2.1.12)</td>
<td>339</td>
</tr>
<tr>
<td>Phosphoglycerate kinase (EC 2.7.2.3)</td>
<td>393</td>
</tr>
<tr>
<td>Phosphoglycerate mutase (EC 5.4.2.12)</td>
<td>552</td>
</tr>
<tr>
<td>Enolase (EC 4.2.1.11)</td>
<td>471</td>
</tr>
<tr>
<td>Pyruvate kinase (EC 2.7.1.40)</td>
<td>500</td>
</tr>
<tr>
<td>Pyruvate formate-lyase (EC 2.3.1.54)</td>
<td>117</td>
</tr>
<tr>
<td>Phosphate acetyltransferase (EC 2.3.1.8)</td>
<td>225</td>
</tr>
<tr>
<td>Acetate kinase (EC 2.7.2.1)</td>
<td>392</td>
</tr>
<tr>
<td>Aldehyde dehydrogenase (EC 1.2.1.3)</td>
<td>494</td>
</tr>
<tr>
<td>Alcohol dehydrogenase (EC 1.1.1.1)</td>
<td>287</td>
</tr>
<tr>
<td>Glycogen phosphorylase (EC 2.4.1.1)</td>
<td>819</td>
</tr>
<tr>
<td>Glycogen synthase (EC 2.4.1.21)</td>
<td>490</td>
</tr>
<tr>
<td>L-lactate dehydrogenase (EC 1.1.1.27)</td>
<td>316</td>
</tr>
<tr>
<td>D-lactate dehydrogenase (EC 1.1.1.28)</td>
<td>324</td>
</tr>
<tr>
<td>Formate dehydrogenase (EC 1.17.1.9)</td>
<td>945</td>
</tr>
<tr>
<td>1,4-α-glucan branching enzyme (EC 2.4.1.18)</td>
<td>198</td>
</tr>
<tr>
<td>Phosphogluco mutase (EC 5.4.2.2)</td>
<td>560</td>
</tr>
</tbody>
</table>

# aa: length in amino acids
Table 2.8. Number of coding sequences (CDS) assigned to predicted amino acid metabolism-related functions in the contig sets from Trials 2, 3 and 4.

<table>
<thead>
<tr>
<th>Enzymes and Transporters</th>
<th>Trial 2 (CDS)</th>
<th>Trial 3 (CDS)</th>
<th>Trial 4 (CDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteases and peptidases</td>
<td>28</td>
<td>36</td>
<td>16</td>
</tr>
<tr>
<td>Branched chain amino acid transporters</td>
<td>24 (4)#</td>
<td>5 (1)#</td>
<td>14 (3)#</td>
</tr>
<tr>
<td>Oligopeptide transporters</td>
<td>6 (2)#</td>
<td>42 (8)#</td>
<td>35 (7)#</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (EC 1.4.1.2)</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Glutamate synthase (EC 1.4.1.13)</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

#Number in parenthesis indicates the number of potential multi-subunit transporter complexes. CDS found sequentially and in close proximity on the same contig were considered to potentially be subunits of a common multi-protein complex.
References


Communities Illustrate Enrichment of Prevotella and Megasphaera in Indian Subjects.


limited number of metabolic pathways between abundant bacteria. Sci Rep, 8(1), 10504. doi:10.1038/s41598-018-28827-7


CHAPTER 3

IDENTIFICATION AND GENOMIC CHARACTERIZATION OF CANDIDATE LACTATE UTILIZING BACTERIA FROM THE RUMEN OF BEEF CATTLE
ABSTRACT

Grain overload in ruminants results in increased production and accumulation of lactate in the rumen. While lactate utilizing bacteria are involved in lactate metabolism, high lactate levels can result in death of these lactate metabolizing bacteria. Therefore, the main objective of this study was to identify previously uncharacterized bacteria involved in metabolizing lactate using a batch culture approach, then predict their functional potential using metagenomics. Three experimental trials, each performed with a different rumen fluid donor, designated as Trial 1, Trial 2, and Trial 3, were conducted during this study. Four replicate cultures, each consisting of approximately 2.2 L of rumen fluid per laboratory-scale bioreactors were used for the study. Three of the replicate cultures were supplemented with lactate (20 ml/L, Sodium DL-lactate solution, Sigma), while the remaining culture was not supplemented with any substrate (Control). Samples collected on days 7 and 14 were analyzed to identify potential lactate utilizers as highly enriched OTUs. Seven species-level OTUs were enriched, which were affiliated to the genera Coprococcus, Butyribrio, Clostridium, Anaerococcus, and Anaerosalibacter. One OTU affiliated to Butyribrio fibrisolvens (LacT2_D14_1_15251) and one OTU affiliated to Anaerococcus prevotii (LT3.7-00154) were then selected for shotgun metagenomics analysis. Annotation using RAST revealed genome-derived coding sequences that could be assembled into predicted pathways. Notably, enzymes that were predicted to be involved in lactate metabolism were identified. Further characterizing of these lactate utilizers (OTUs) may provide additional insights towards a better understanding of lactate metabolism and also the events that impact the onset of acidosis.
1. Introduction

The rumen is a complex and diverse microbial eco-system that is responsible for fermenting feedstuffs into SCFAs and producing microbial proteins (Bryant, 1959; Robert E. Hungate, 1966). Since approximately 70% of the energy of the host is provided by SCFAs, ruminal microbial communities are essential contributors to ruminant productivity (Bergman, 1990).

With the aim of increasing their production, feedlot beef and dairy cattle are commonly fed with grain-rich diets. However, grain overload is associated with excessive production and accumulation of ruminal SCFAs, resulting in a pH drop in the rumen (J. Hernández, J. L. Benedito, A. Abuelo, & C. Castillo, 2014). This condition, known as acidosis, can be categorized as either acute or subacute. A rumen pH below 5.0, characterized by increased lactic acid production, is considered acute acidosis, whereas a rumen pH between 5.0 and 5.6 is characterized by increased accumulation of SCFAs and is regarded as subacute acidosis (Krause & Oetzel, 2006; Nagaraja & Titgemeyer, 2007; Owens, Secrist, Hill, & Gill, 1998). Increased lactic acid production and decreased lactic acid utilization during acute ruminal acidosis is due to the establishment of acid tolerant bacteria as well as to a decrease in the abundance of lactic acid fermenting bacteria (Therion, Kistner, & Kornelius, 1982) (Russell & Hino, 1985) (Mackie & Gilchrist, 1979a).

Various studies have reported that *Streptococcus bovis* and *Lactobacillus* spp are dominant lactic acid-producing bacteria, while *Megasphaera elsdenii* and *Selenomonas ruminantium subsp. lactilytica* are potent lactic acid utilizing bacteria (Mackie & Gilchrist, 1979b) (Ogunade, Pech-Cervantes, & Schweickart, 2019) (R. E. Hungate,
Dougherty, Bryant, & Cello, 1952) (Mackenzie, 1967) (Chen et al., 2019). In animal nutritional models, experimental induction of SARA have been associated with reduced bacterial diversity and richness (Plaizier et al., 2017) (Plaizier, Li, Tun, & Khafipour, 2016) (Plaizier, 2004) (E. Khafipour, Krause, & Plaizier, 2009) (Mao, 2017). A study by McCann et al. (2016) reported that cows with induced SARA had a higher relative abundance of Bacteroidetes in the liquid fraction, whereas an increase in the relative abundance of Firmicutes was observed in the solid fraction of the rumen samples (McCann et al., 2016). A study by Khafipour et al. (2009) identified increased relative abundance of Bacteroidetes when cows were induced with SARA by feeding alfalfa pellets when compared with the grain diet induced SARA (Ehsan Khafipour, Li, Plaizier, & Krause, 2009). On their part, a study by Ogunade et al. (2019) investigated the metabolic potential of the rumen microbiome during SARA. It was reported that 68 genes could be mapped to carbohydrate, amino acid, energy, vitamin and co-factor metabolism pathways, and that bacterial biofilm formation pathways were highly expressed in the SARA-induced beef cows (Ogunade et al., 2019).

Considering that typically only a fraction of microorganisms identified in rumen samples correspond to valid species, it is generally acknowledged that the vast majority of ruminal symbionts remain to be characterized (Creevey, Kelly, Henderson, & Leahy, 2014) (Huws et al., 2018). Based on this assessment, it was hypothesized that currently unknown ruminal microorganisms include lactic acid fermenting bacteria that have yet to be identified. Using a batch-culture system, candidate lactate utilizers were enriched from beef cow rumen fluid, then characterized as Operational Taxonomic Units (OTUs) using a 16S rRNA-based approach. Metagenomics analyses were conducted to assess the
metabolic potential of candidate ruminal bacterial species, which revealed coding sequences for enzymes that participate in pathways predicted to metabolize lactate into SCFAs. Together, the results presented in this chapter suggest that the bacterial species corresponding to these OTUs function as lactate utilizers in the rumen of beef cows.

2. Materials and Methods

2.1 Sample collection and in vitro rumen culture experiments

All animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at South Dakota State University. Fresh rumen fluid was obtained by manually squeezing ruminal digesta collected from fistulated beef cows maintained at the South Dakota State University Cow-Calf Research Facilities. The diet of these animals consisted of pasture hay or haylage during the period when rumen fluid was collected (March 2017 to November 2017). Three experimental trials, each performed with a different rumen fluid donor, designated as Trial 1, Trial 2, and Trial 3, were conducted during this study, with the following procedure followed for each trial. Four replicate cultures, each consisting of approximately 2.2 L of rumen fluid per laboratory-scale bioreactors (Chemglass), were set up within one hour of collection. Three of the replicate cultures were supplemented with lactate (20 ml/L, Sodium DL-lactate solution, sigma), while the remaining culture was not supplemented with any substrate (Control). Treatment and control cultures were maintained under anaerobic conditions at a constant physiological temperature (38°C), with continuous agitation at 150 rpm using a Rushton-style impeller fixed to a stirring shaft built into the bioreactor. Each bioreactor allowed
for approximately 0.8 L of headspace, and was equipped with a flexible plastic tube for exhaust of excess biogas to prevent pressure build-up. Culture samples (approximately 15 mL/sample) were collected on days 7 (D7) and 14 (D14) to assess their respective bacterial composition. Samples from the rumen inoculum (D0) and cultures were stored frozen at -20°C until analyzed.

2.2 Microbial genomic DNA purification and PCR amplification of the 16S rRNA gene

Genomic DNA was extracted from each sample by a repeated bead beating plus column method as previously described (Yu & Morrison, 2004). Briefly, 250µL of rumen sample was lysed in extraction buffer (0.5 M NaCl, 50 mM Tris.HCl, 50Mm EDTA, 4% SDS) by bead beating followed by sequential extraction with 10M ammonium acetate then isopropanol precipitation. Recovered nucleic acids were then purified using the QIAamp DNA Stool Kit (QIAGEN, Hilden, Germany) following the manufacturer’s recommended protocol. PCR was performed using the Phusion Taq DNA polymerase (ThermoFischer Scientific, Waltham, MA, USA) on a 2720 Thermo Cycler (ThermoFischer Scientific), with the 27F (Edwards, Rogall, Blocker, Emde, & Bottger, 1989) and 519R (Lane et al., 1985) primers to target the V1-V3 regions of the 16S rRNA gene. PCR reactions consisted of a ‘hot start’ (98°C, 3min), followed by 35 consecutive cycles of denaturation (98°C, 30 sec), annealing (50°C, 30 sec), and elongation (72°C, 30 sec), then by a final elongation period (72°C, 10 min). Quality of PCR amplicons (expected approximate length of 500 bp) was determined by agarose gel electrophoresis, with recovery of PCR-generated DNA using the QiaexII Gel extraction kit (Qiagen, Hilden, Germany). Gel purified amplicons from each sample were then submitted as
template for high throughput sequencing using an Illumina MiSeq(2×300) platform (Molecular Research DNA, Shallowater, TX).

2.3 Bioinformatic analysis for 16S rRNA gene-based composition analysis

Unless specified, datasets were analyzed using custom written Perl scripts (available upon request). Raw bacterial 16S rRNA gene V1-V3 amplicon sequences were provided by Molecular Research DNA as contigs assembled from overlapping MiSeq 2x300 paired-end reads from the same flow cell clusters. Reads were then selected to meet the following criteria: presence of both intact 27F (forward) and 519R (reverse) primer nucleotide sequences, length between 400 and 580 nt, and a minimal quality threshold of no more than 1% of nucleotides with a Phred quality score lower than 15.

Following quality screens, sequence reads were aligned, then clustered into Operational Taxonomic Units (OTUs) at a genetic distance cutoff of 5% sequence dissimilarity (St-Pierre & Wright, 2015). While 3% is the most commonly used clustering cutoff for 16S rRNA, it was originally recommended for full length sequences, and may not be suitable for the analysis of specific sub-regions since nucleotide sequence variability is not constant across the entire length of the 16S rRNA gene. In this context, if 3% is a commonly accepted clustering cutoff for V4 or V4-V5 regions, which are the least variable of the hypervariable regions, then a higher cutoff should be used for the V1-V3 region, since V1 is the most variable region of the 16S rRNA gene. OTUs were screened for DNA sequence artifacts using the following methods. Chimeric sequences were first identified with the chimera.uchime and chimera.slayer commands from the MOTHUR open source software package (Schloss et al., 2009). Secondly, the integrity of the 5’ and 3’ ends of OTUs was evaluated using a database alignment search-based
approach; when compared to their closest match of equal or longer sequence length from the NCBI nt database, as determined by BLAST (Altschul et al., 1997), OTUs with more than five nucleotides missing from the 5’ or 3’ end of their respective alignments were discarded as artifacts. Single read OTUs were subjected to an additional screen, where only sequences that had a perfect or near perfect match to a sequence in the NCBI nt database were kept for analysis, i.e. that the alignment had to span the entire sequence of the OTU, and a maximum of 1% of dissimilar nucleotides was tolerated.

After removal of sequence chimeras and artifacts, RDP Classifier (Qiong Wang, Garrity, Tiedje, & Cole, 2007) and BLAST (Altschul et al., 1997) were used for taxonomic assignment of valid OTUs. The List of Prokaryotic Names with Standing in Nomenclature (LPSN) was also consulted for information on valid species belonging to taxa of interest (Parte, 2014).

2.4 Metagenomic analysis

A select number of samples exhibiting high levels of enrichment for candidate lactate utilizing OTUs were further analyzed using a metagenomics approach. Purified microbial genomic DNA (extracted as described above) was used directly as template for high throughput sequencing using an Illumina Miseq(2×250) platform (Molecular Research DNA, Shallowater, TX). Raw sequence reads of 200 bp in length or longer were selected using custom written Perl scripts for building of genomic contigs using the de novo assembly program ABySS (Simpson et al., 2009). ABySS was run on a Linux-based high-performance computing cluster maintained by the University Networking and Research Computing Group at South Dakota State University, with k-mer length set at 64 and all other parameters used at their respective default values.
ABYSS-assembled contigs of 4000 bp or greater were selected for further analysis by gene annotation. Coding sequences (CDS) were identified and annotated using RAST (Rapid Annotations using Subsystems Technology) (Aziz et al., 2008), with assignment of CDS to metabolic pathways performed using KEGG pathways as a model reference (Kanehisa & Goto, 2000). Further analysis of select CDS was performed against the NCBI nr database using BLASTp (Altschul et al., 1997) to identify their respective closest homologs.

3. Results

3.1 Comparative analysis of bacterial communities from donor rumen fluid

Analysis of the rumen fluid inocula before culturing with lactate revealed that bacteria affiliated to the phyla Bacteroidetes and Firmicutes were overall the most abundant, representing 45.4-58.4% and 30-47% of sequences reads, respectively (Figure 3.1). Prevotellaceae was the predominant Bacteroidetes family (17.1-31.3%), while unclassified Clostridiales was the most highly represented group of Firmicutes at the family level (28.2-34.8%, respectively).

3.2 Identification of candidate bacterial lactate utilizers from rumen fluid

When compared to their respective culture controls without substrate, seven distinct OTUs were found to be the most enriched bacteria in lactate-supplemented rumen fluid cultures (Table 3.1).
From the Trial 1 donor, OTU LacT1_D7_3-00877 was in highest abundance in lactate supplemented cultures at day 7 with a relative abundance of 0.16%-22%. The closest valid relative to LacT1_D7_3-00877 was *Coprococcus catus*, a well characterized lactate utilizer (Holdeman & Moore, 1974a), with a sequence identity of 93% based on 16S rRNA gene-based analyses. OTUs LacT2_D14_1-15251 and LacT2_D14_2-12405 were identified as candidate lactate utilizers from Trial 2. These OTUs were in highest abundance in lactate supplemented cultures at day 14 with a relative abundance of 0.6%-27% and 0.05%-13%, respectively. The closest valid relative to LacT2_D14_1-15251 and LacT2_D14_2-12405 were respectively *Butyrivibrio fibrisolvens* (Islam et al., 2019) and *Clostridium acetiредucens* (Örlygsson, Krooneman, Collins, Pascual, & Gottschal, 1996), with a sequence identity 89% and 99%, based on 16S rRNA gene-based analyses. OTUs LT3.7-00154, LT3.7-01259, LT3.7-08849, and LT3.8-08697 were identified as candidate lactate utilizers from Trial 3. Three of the OTUs (LT3.7-00154, LT3.7-01259, and LT3.7-08849) were in highest abundance in lactate supplemented cultures at day 14, with a relative abundance of 0.01%-39%, 0.05%-16%, and 0.01%-13%, whereas OTU LT3.8-08697 was in highest abundance in lactate supplemented cultures at day 7, with a relative abundance of 0.03%-11%. The closest valid relative to OTUs LT3.7-00154, LT3.7-01259, LT3.7-08849, and LT3.8-08697 were respectively *Anaerococcus prevotii* (Labutti et al., 2009), *Anaerosalibacter bizertensis* (Rezgui et al., 2012), *Anaerococcus tetradius* (Cibis, Gneipel, & Konig, 2016), and *Clostridium cochlearium* (Schwab, Rago, Koch, & Harnisch, 2019) with a sequence identity of 96%, 99%, 93%, and 99%.

### 3.3 Exploring the metabolic potential of OTUs identified as candidate lactate utilizers

...
We postulated that sequence datasets generated from samples with high abundance of an OTU would have higher representation of genomic sequences from this OTU. Consequently, higher sequence representation or coverage would then be favorable for the assembly of contigs that would be of sufficient length to allow for prediction of full-length coding sequences and improved accuracy of gene annotation. From the seven OTUs that were enriched in the presence of lactate, we focused our attention on two highly enriched OTUs affiliated to Butyrivibrio fibrisolvens (LacT2_D14_1_15251) and Anaerococcus prevotii (LT3.7-00154).

The most highly enriched samples for LacT2_D14_1_15251 (Trial 2: 27%) and LT3.7-00154 (Trial 3: 39%), were then selected for shotgun sequencing, which generated a total of 9.4 X 10^6 and 9.1 X 10^6 paired end reads, respectively. After de novo assembly, contigs were filtered for a minimal length of 4,000 nt, which resulted in the selection of 324 (Trial 2) and 682 (Trial 3) contigs for further analysis. The maximum contig length from each sample was found to be 38,502 nt (Trial 2) and 24,827 nt (Trial 3). The RAST online tool was then used for predicting amino acid coding regions (CDS) from the selected contigs, as well as for performing gene annotation. For the Trial 2 contig set, 860 of 2,759 CDS were annotated successfully, compared to 741 of 5,629 CDS for Trial 3.

Predicted metabolic functional categories from gene annotation included carbohydrates, proteins, amino acids and derivatives, DNA and RNA, phosphorus, potassium and sulfur, regulation and cell signaling, as well as stress response (Figure 3.2). Based on the conditions of the culturing assay, our primary interest was to identify genes involved in carbohydrate metabolism; 80 CDS were assigned to this function from the Trial 2 contig set, compared to 104 CDS for the Trial 3 dataset.
To gain further insight, data mining for specific functions within carbohydrate metabolism was performed, resulting in the identification of enzymes predicted to be associated with metabolic activities of lactate utilizing bacteria (Figures 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, and 3.9).

Coding sequences for lactate dehydrogenase (EC 1.1.2.4, EC 1.1.1.27, EC 1.1.1.28), which can catalyze the conversion of lactate to pyruvate, were identified for LacT2_D14_1_15251 and for LT3.7-00154 (Figure 3.3 and 3.6).

For both OTUs, enzymes involved in metabolizing pyruvate into SCFAs were also identified (Figure 3.3 and 3.6). In trial 2 and 3, we identified enzymes such as pyruvate dehydrogenase (EC 1.2.4.1), the dihydrolipoamide acetyltransferase component of the pyruvate dehydrogenase complex (EC 2.3.1.12), as well as pyruvate-formate lyase (EC 2.3.1.54) which converts pyruvate into acetyl-CoA. From the trial 2 dataset, acetyl-CoA could then be metabolized into acetate by succinyl-CoA:acetate-CoA-transferase (EC 2.8.3.18), while from the trial 3 dataset, acetyl-CoA could be metabolized sequentially into acetyl phosphate (phosphate acetyltransferase, EC 2.3.1.8) then acetate (acetate kinase, EC 2.7.2.1). Acetaldehyde dehydrogenase (EC 1.2.1.10) and alcohol dehydrogenases (EC 1.1.1.1) found in both trial 2 and 3 could catalyze the conversion of acetyl-CoA to acetaldehyde and ethanol. Enzymes involved in butanoate metabolism and the citrate cycle were also identified in LacT2_D14_1_15251 affiliated contigs (Figure 3.4 and 3.5).

All enzymes of the glycolysis pathway (Figure 3.7) and enzymes involved in glycogen synthesis and utilization (Figure 3.8) were also identified in LT3.7-00154 affiliated contigs. Enzymes predicted to be involved in this function included glycogen
synthase (EC 2.4.1.21), which converts ADP-glucose to amylose, as well as 1,4-alpha-glucan branching enzyme (EC 2.4.1.18), which generates glycogen from amylose. Glucose stored as glycogen would then be predicted to be metabolized through the sequential actions of glycogen phosphorylase (EC 2.4.1.1) and phosphoglucomutase (EC 5.4.2.2) which would generate glucose-1-phosphate and glucose-6-phosphate, respectively. The latter could then enter glycolysis to be metabolized into pyruvate and SCFAs.

Contig sets from LacT2_D14_1_15251 and LT3.7-00154 enriched samples included enzymes (EC 1.4.1.13, EC 6.3.5.4, EC 6.3.1.2, EC 1.4.1.4, EC 1.4.1.2) that were involved in amino acid metabolism (Table 3.4 and 3.5). Monocarboxylate and ammonium transporters were identified in contig sets from LacT2_D14_1_15251 and LT3.7-00154 enriched samples and were involved in lactate and amino acid transportation (Table 3.6).

4. Discussion

Acidosis is one of the most commonly seen clinical disorders in intensive management systems when ruminants are fed with high amounts of grain. This condition has a significant impact on rumen microbial populations and their cell densities, further leading to negative effects on animal health and productivity (Joaquín Hernández et al., 2014). Lactate is one of the main contributors to acidosis, as increases in its production and its accumulation can have a severe impact on ruminal pH (Mackie & Gilchrist, 1979b). By fermenting lactate, lactate utilizing bacteria can counter this effect, and therefore contribute to regulating lactate levels in the rumen.
However, one critical gap in our knowledge is that most lactate metabolizing bacteria likely remain to be identified and characterized, as an estimated 95% of ruminal bacterial species have yet to be assigned a function (Creevey et al., 2014). In this context, the primary objective of the research presented in this chapter was to identify uncharacterized ruminal bacteria that can metabolize lactate, then assess their metabolic potential using a metagenomic approach. The bacterial composition of the three rumen samples prior to culturing in the presence of lactate (D0), in which Bacteroidetes and Firmicutes were the most abundant phyla and Prevotellaceae were overall the most predominant family, was consistent with previously reported studies (Pitta et al., 2014) (Lima et al., 2015) (Opdahl, Gonda, & St-Pierre, 2018) (Jami & Mizrahi, 2012) (Jami, White, & Mizrahi, 2014) (Jewell, McCormick, Odt, Weimer, & Suen, 2015). From these complex communities, seven main OTUs were found to be in higher relative abundance in ruminal fluid cultures when lactate was provided as the only supplemented substrate.

Enrichment for OTU LacT1_D7_3-00877, predicted to be a strain of Coprococcus catus based on 16S rRNA gene sequence comparisons, was enriched on day-7 with a relative abundance of 22% in one of the lactate supplemented culture. Coprococcus catus has been reported as a key lactate fermenter from various studies (Holdeman & Moore, 1974b) (Reichardt et al., 2014; Shabat et al., 2016).

From trial 2, LacT2_D14_1-15251 was found to be enriched in a single culture with an abundance of 27%, and it was most closely related to Butyrivibrio fibrisolvens. A Study by Ghali et al. (2011) on the foregut of feral camels described three isolates closely related to Butyrivibrio fibrisolvens NCDO 2398 strain, which have been reported to be involved in lactate metabolism, based on culturing with lactic acid medium (LAM) and
De Man, Rogosa, Sharpe medium (MRS) medium (Ghali, Scott, Alhadrami, & Al Jassim, 2011). Another study reported that certain strains of *Butyrivibrio fibrisolvens* have proteolytic activity in the rumen (Cotta & Hespell, 1986). Our observations are then consistent with these previously published results, as they indicate that OTU LacT2_D14_1-15251 could function as a lactate utilizer. The other OTU enriched in trial 2, LacT2_D14_2-12405, with its highest abundance in a single culture at 13%, was assigned to *Clostridium acetireducens*. A study by Orlygsson et al. (1996) describes the isolation of *Clostridium acetireducens* from an anerobic bio-reactor fed on waste from potato starch. While it was not reported as a lactate-utilizer, this anerobic bacteria was found to be capable of deaminating amino acids by using acetate as an electron acceptor while producing butyrate as the reduced end product (Örlygsson et al., 1996).

Four OTUs from trial 3, LT3.7-00154, LT3.7-01259, LT3.7-08849, and LT3.8-08697, were assigned to three different lineages. OTUs LT3.7-00154 and LT3.7-08849 were found to be affiliated to the genus *Anaerococcus*, with members having been isolated from human skin, oral cavity, gastro-intestinal environment, vaginal discharges and ovarian abscess (Ezaki, Yamamoto, Ninomiya, Suzuki, & Yabuuchi, 1983) (Cibis et al., 2016). While they have been reported to ferment different sugars (Ezaki et al., 1983; Murdoch, 1998), no studies have reported the potential role of these species as lactate utilizers. For its part, OTU LT3.7-01259 was most closely affiliated to *Anaerosalibacter bizertensis*, a bacterial species isolated from sludge, which has the ability to ferment different carbon sources such as glucose, succinate, and pyruvate (Rezgui et al., 2012). However, no studies have documented the role of *Anaerosalibacter bizertensis* as a lactate utilizer. Finally, OTU LT3.8-08697 was assigned to *Clostridium cochlearium*, a
bacterial species originally isolated from the mouse gut microbiome (Schwab et al., 2019) and from a slaughter house (Bakhtiyari et al., 2018). Known metabolic activities include fermentation of glucose or glutamate to acetate, butyrate, and propionate (Buckel & Barker, 1974; Cibis et al., 2016), but no reported ability to metabolize lactate.

In light of the limitations of 16S rRNA-based composition analyses to provide functional insights (Creevey et al., 2014) (Deusch, Tilocca, Camarinha-Silva, & Seifert, 2015), a metagenomics analysis was performed on representative high enrichment samples to get an assessment of the metabolic potential of OTUs LacT2_D14_1-15251 and LT3.7-00154. Consistent with lactate utilization, both contig sets included coding sequences for monocarboxylate transporters, which could be involved in lactate transport across biological membranes (Thauer, 1988) (Pösö, 2002), as well as for lactate dehydrogenases (EC 1.1.2.4, EC 1.1.1.28, and EC 1.1.1.27), which would catalyze the conversion of lactate to pyruvate (Garvie, 1980) (J. Hernández, J. Benedito, A. Abuelo, & C. Castillo, 2014) (Allison, O'Donnell, Hoey, & Fewson, 1985; Goffin, Lorquet, Kleerebezem, & Hols, 2004). A coding sequence for D-lactate ferricytochrome c oxidoreductase (EC 1.1.2.4), known to be present in Saccharomyces cerevisiae, Pseudomonas stutzeri, and Acinetobacter (Lodi & Ferrero, 1993) (Pajot & Claisse, 1974) (C. Ma et al., 2007) (Hao et al., 2007) (C. Q. Ma et al., 2004), was predicted from the contigs assembled from the sample enriched for LacT2_D14_1-15251. Coding sequences for NAD-dependent lactate dehydrogenase enzymes (EC 1.1.1.27 and EC 1.1.1.28), which were identified in the contigs associated with LT3.7-00154, are involved in catalyzing the conversion of lactate to pyruvate. Besides Anaerococcus prevotii,
*Lactobacillus plantarum* can also express this enzyme under aerobic conditions (Goffin et al., 2004).

A number of different enzymes responsible for metabolizing pyruvate into SCFAs were also identified. CDS for enzymes such as pyruvate dehydrogenase complex (EC 1.2.4.1, EC 2.3.1.12), pyruvate formate-lyase (EC 2.3.1.54), acetaldehyde dehydrogenase (EC 1.2.1.10), and alcohol dehydrogenase (EC 1.1.1.1) were identified in both contig sets analyzed, indicating that both LacT2_D14_1-15251 and LT3.7-00154 had the potential to metabolize pyruvate to SCFAs. Pyruvate dehydrogenase, pyruvate formate-lyase, aldehyde dehydrogenase, and alcohol dehydrogenase enzymes can all be encoded in the same bacterial genome, as shown for *Escherichia coli* (Clark, 1989) (Qingzhao Wang, Ou, Kim, Ingram, & Shanmugam, 2010) (Neidhardt & Curtiss, 1996). In addition, CDS for all enzymes of the citrate cycle were identified from the contig set analyzed from trial 2. Previous studies have reported the presence of citrate cycle enzymes in aerobic bacteria (Alteri, Himpsl, Engstrom, & Mobley, 2012; Galushko & Schink, 2000; Thauer, 1988). CDS for all enzymes of the glycolysis pathway were identified in the contig set for LT3.7-00154, indicating that this OTU had the potential to metabolize glucose into pyruvate (Rodríguez, Sosa, & Rodríguez, 2007). Enzymes that could synthesize glycogen from glucose were also identified from the trial 3 contig set.

Amino acid metabolism is also essential for the proliferation of ruminal microorganism. Based on the predicted enzymatic capabilities from the contig sets analyzed in this study, both enriched OTUs appeared to have enzymes such as glutamate synthase (EC 1.4.1.13), glutamate dehydrogenase (EC 1.4.1.2), glutamine synthetase type I (EC 6.3.1.2), glutamate dehydrogenase (EC 1.4.1.4), asparagine synthetase (EC 6.3.5.4).
and ammonium transporters for the synthesis of amino acids. The presence of ammonium transporters indicates that OTUs LacT2_D14_1-15251 and LT3.7-00154 could use ammonia from the extracellular environment for the synthesis of amino acids (Pengpeng & Tan, 2013) (Hackette, Skye, Burton, & Segel, 1970) (Marini, Vissers, Urrestarazu, & André, 1994) (Phibbs & Bernlohr, 1971) (Erfle et al., 1977).

5. Conclusion

The batch-culture approach used in this study allowed for the identification of previously uncharacterized bacteria as candidate lactate utilizers from the rumen of beef cows. Based on 16S rRNA gene sequence analysis, seven ruminal bacterial OTUs enriched with lactate as the only provided substrate were found to be affiliated to the genera Coprococcus, Butyrivibrio, Clostridium, Anaerococcus, and Anaerosalibacter. Further investigations of the Butyrivibrio and Anaerococcus related OTUs using metagenomics revealed genome-derived coding sequences that could be assembled into predicted pathways that would allow for fermentation of lactate into SCFAs. Furthermore, these different OTUs shared other metabolic functions, such as the ability to store glucose as glycogen, as well as to utilize extracellular ammonia for the synthesis of amino acids. Together, these characteristics could help define members of a particular subgroup of ruminal lactate utilizers, ultimately contributing to a better understanding of metabolic functions and events that impact the onset of SARA and the incidence of acidosis.
Figure 3.1. Taxonomic composition of rumen samples used as inocula, prior to setting up batch culture enrichment experiments with lactate. Each experimental trial (T1, T2, T3) is represented on the horizontal axis, with the relative abundance (%) of each major taxonomic group shown on the vertical axis. Families belonging to the same phylum are represented by different shades of the same color: Firmicutes (green), Bacteroidetes (blue), and other families (red).
Figure 3.2. RAST gene annotation analysis of contig sets generated from enrichment Trials 2 (blue) and 3 (orange). The relative abundance (%) of coding sequences for each subsystem level category is shown.
A) Figure 3.3. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomics analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to *Butyrivibrio fibrosolvens*). A) lactate and pyruvate metabolism. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.2 for corresponding enzyme names.
Figure 3.4. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#2 (OUT LacT2_D14_1-15251, affiliated to *Butyrivibrio fibrosolvens*). B) butanoate metabolism. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.2 for corresponding enzyme names.
Figure 3.5. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to *Butyribrio fibrosolvens*). C) citrate cycle. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.2 for corresponding enzyme names.
Figure 3.6. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#3 (OTU LT 3.7-00154, affiliated to Anaerococcus prevotii). A) lactate and pyruvate metabolism. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.3 for corresponding enzyme names.
Figure 3.7. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of predicted metabolic pathways from the metagenomic analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*). B) glycolysis pathway. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.3 for corresponding enzyme names.
Figure 3.8. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of precited metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*). C) glycogen synthesis and mobilization. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.3 for corresponding enzyme names.
Figure 3.9. KEGG (Kyoto Encyclopedia of Genes and Genomes) annotation of precited metabolic pathways from the metagenomics analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*). D) pentose phosphate pathway. Each enzyme is represented by its Enzyme Commission (EC) number, which is based on the chemical reaction it catalyzes. Please consult Table 3.3 for corresponding enzyme names.
Table 3.1. Enrichment of OTUs in response to lactate. Abundance is presented as relative abundance (%) of total non-chimeric reads. D0 (prior to lactate enrichment), C-D7 and C-D14 (Control culture day-7 and day-14), L-D7 and L-D14 (Lactate culture day-7 and day-14).

<table>
<thead>
<tr>
<th>Trial</th>
<th>OTU</th>
<th>D0</th>
<th>C-D7</th>
<th>L-D7*</th>
<th>C-D14</th>
<th>L-D14*</th>
<th>Closest relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial 1</td>
<td>LacT1_D7-3-00877</td>
<td>6.7x10^{-05}</td>
<td>4.7x10^{-05}</td>
<td>11</td>
<td>0</td>
<td>4</td>
<td>C. catus (93%)</td>
</tr>
<tr>
<td>Trial 2</td>
<td>LacT2_D14_1-15251</td>
<td>0</td>
<td>0</td>
<td>0.8</td>
<td>0</td>
<td>11</td>
<td>B. fibrosolvens (89%)</td>
</tr>
<tr>
<td></td>
<td>LacT2_D14_2-12405</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
<td>4</td>
<td>C. acetireducens (99%)</td>
</tr>
<tr>
<td>Trial 3</td>
<td>LT3.7-00154</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
<td>0.03</td>
<td>13</td>
<td>A. prevotii (96%)</td>
</tr>
<tr>
<td></td>
<td>LT3.7-01259</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>6.8x10^{-05}</td>
<td>6</td>
<td>A. bizertensis (99%)</td>
</tr>
<tr>
<td></td>
<td>LT3.7-08849</td>
<td>3.5x10^{-05}</td>
<td>0.02</td>
<td>0.06</td>
<td>0.02</td>
<td>4</td>
<td>A. tetradius (93%)</td>
</tr>
<tr>
<td></td>
<td>LT3.8-08697</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.03</td>
<td>6.8x10^{-05}</td>
<td>0.02</td>
<td>C. cochlearium (99%)</td>
</tr>
</tbody>
</table>
Table 3.2. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to \textit{Butyrivibrio fibrisolvens}).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa#</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-Lactate dehydrogenase, cytochrome c-dependent (EC 1.1.2.4)</td>
<td>482</td>
<td>0</td>
</tr>
<tr>
<td>Pyruvate dehydrogenase (EC 1.2.4.1)</td>
<td>258</td>
<td>3x10^{-05}</td>
</tr>
<tr>
<td>Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)</td>
<td>564</td>
<td>0</td>
</tr>
<tr>
<td>Pyruvate kinase (EC 2.7.1.40)</td>
<td>810</td>
<td>0</td>
</tr>
<tr>
<td>Acetaldehyde dehydrogenase (EC 1.2.1.10)</td>
<td>319</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol dehydrogenase (EC 1.1.1.1)</td>
<td>273</td>
<td>5x10^{-05}</td>
</tr>
<tr>
<td>Malate synthase (EC 2.3.3.9)</td>
<td>725</td>
<td>0</td>
</tr>
<tr>
<td>Succinyl-CoA: acetate CoA-transferase (EC 2.8.3.18)</td>
<td>503</td>
<td>0</td>
</tr>
<tr>
<td>3-Hydroxybutyryl-CoA epimerase (EC 5.1.2.3)</td>
<td>317</td>
<td>1x10^{-5}</td>
</tr>
<tr>
<td>3-Hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157)</td>
<td>317</td>
<td>2x10^{14}</td>
</tr>
<tr>
<td>Acetyl-CoA acetyltransferase (EC 2.3.1.9)</td>
<td>405</td>
<td>0</td>
</tr>
<tr>
<td>Enoyl-CoA hydratase (EC 4.2.1.17)</td>
<td>272</td>
<td>1x10^{-12}</td>
</tr>
<tr>
<td>Butyryl-CoA dehydrogenase (EC 1.3.8.1)</td>
<td>385</td>
<td>0</td>
</tr>
<tr>
<td>Fumarate hydratase class I (EC 4.2.1.2)</td>
<td>509</td>
<td>0</td>
</tr>
<tr>
<td>Malate dehydrogenase (EC 1.1.1.37)</td>
<td>346</td>
<td>0</td>
</tr>
<tr>
<td>Citrate synthase (si) (EC 2.3.3.1)</td>
<td>436</td>
<td>0</td>
</tr>
<tr>
<td>Aconitate hydratase 2 (EC 4.2.1.3)</td>
<td>863</td>
<td>0</td>
</tr>
<tr>
<td>Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)</td>
<td>418</td>
<td>0</td>
</tr>
<tr>
<td>2-Oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)</td>
<td>957</td>
<td>0</td>
</tr>
<tr>
<td>Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)</td>
<td>293</td>
<td>0</td>
</tr>
<tr>
<td>Succinate dehydrogenase flavoprotein subunit (EC 1.3.5.1)</td>
<td>568</td>
<td>0</td>
</tr>
</tbody>
</table>

# aa: length in amino acids
Table 3.3. Carbohydrate-utilizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa(^a)</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-Lactate dehydrogenase (EC 1.1.1.27)</td>
<td>315</td>
<td>0</td>
</tr>
<tr>
<td>Nickle-dependent lactate racemase (EC 5.1.2.1)</td>
<td>430</td>
<td>0</td>
</tr>
<tr>
<td>D-lactate dehydrogenase (EC 1.1.1.28)</td>
<td>405</td>
<td>0</td>
</tr>
<tr>
<td>Pyruvate formate lyase (EC 2.3.1.54)</td>
<td>692</td>
<td>0</td>
</tr>
<tr>
<td>Phosphate acetyltransferase (EC 2.3.1.8)</td>
<td>329</td>
<td>0</td>
</tr>
<tr>
<td>Acetate kinase (EC 2.7.2.1)</td>
<td>397</td>
<td>0</td>
</tr>
<tr>
<td>Acetaldehyde dehydrogenase (EC 1.2.1.10)</td>
<td>384</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol dehydrogenase (EC 1.1.1.1)</td>
<td>317</td>
<td>0.13</td>
</tr>
<tr>
<td>Pyruvate phosphate dikinase (EC 2.7.9.1)</td>
<td>874</td>
<td>0</td>
</tr>
<tr>
<td>Pyruvate kinase (EC 2.7.1.40)</td>
<td>592</td>
<td>0</td>
</tr>
<tr>
<td>Phosphogluco mutase (EC 5.4.2.2)</td>
<td>525</td>
<td>0</td>
</tr>
<tr>
<td>Glucose-6-phosphate isomerase (EC 5.3.1.9)</td>
<td>504</td>
<td>0</td>
</tr>
<tr>
<td>6-phosphofructokinase (EC 2.7.1.11)</td>
<td>669</td>
<td>0</td>
</tr>
<tr>
<td>Fructose-bisphosphate aldolase (EC 4.1.2.13)</td>
<td>319</td>
<td>0</td>
</tr>
<tr>
<td>Glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)</td>
<td>287</td>
<td>7x10(^{-17})</td>
</tr>
<tr>
<td>Phosphoglycerate kinase (EC 2.7.2.3)</td>
<td>334</td>
<td>0</td>
</tr>
<tr>
<td>Phosphoglycerate mutase (EC 5.4.2.12)</td>
<td>564</td>
<td>0</td>
</tr>
<tr>
<td>Enolase (EC 4.2.1.11)</td>
<td>397</td>
<td>0</td>
</tr>
<tr>
<td>Glucose-1-phosphate adenyltransferase (EC 2.7.7.27)</td>
<td>412</td>
<td>0</td>
</tr>
<tr>
<td>Glycogen synthase (EC 2.4.1.21)</td>
<td>510</td>
<td>0</td>
</tr>
<tr>
<td>1,4-α-glucan (glycogen) branching enzyme (EC 2.4.1.18)</td>
<td>663</td>
<td>0</td>
</tr>
<tr>
<td>Glycogen phosphorylase (EC 2.4.1.1)</td>
<td>751</td>
<td>0</td>
</tr>
<tr>
<td>Transketolase (EC 2.2.1.1)</td>
<td>111</td>
<td>4x10(^{-4})</td>
</tr>
<tr>
<td>Ribulose-phosphate-3-epimerase (EC 5.1.3.1)</td>
<td>136</td>
<td>1x10(^{-8})</td>
</tr>
<tr>
<td>Ribose-5-phosphate isomerase (EC 5.3.1.6)</td>
<td>154</td>
<td>7x10(^{-7})</td>
</tr>
<tr>
<td>Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)</td>
<td>327</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) aa: length in amino acids
Table 3.4. Amino acid metabolizing enzymes identified from the gene annotation analysis of contigs generated from Trial#2 (OTU LacT2_D14_1-15251, affiliated to *Butyrivibrio fibrisolvens*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa#</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glutamate synthase (NADPH) large chain (EC 1.4.1.13)</td>
<td>1050</td>
<td>0</td>
</tr>
<tr>
<td>Aspargine synthetase (EC 6.3.5.4)</td>
<td>631</td>
<td>0</td>
</tr>
<tr>
<td>Glutamine synthetase type I (EC 6.3.1.2)</td>
<td>481</td>
<td>0</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (NADP+) (EC 1.4.1.4)</td>
<td>431</td>
<td>0</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (EC 1.4.1.2)</td>
<td>450</td>
<td>0</td>
</tr>
</tbody>
</table>

# aa: length in amino acids
Table 3.5. Amino acid metabolizing enzymes identified from the gene annotation analysis of contigs generated from Trial#3 (OTU LT3.7-00154, affiliated to *Anaerococcus prevotii*).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>aa</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glutamate synthase (NADPH) large chain (EC 1.4.1.13)</td>
<td>609</td>
<td>0</td>
</tr>
<tr>
<td>Aspargine synthetase (EC 6.3.5.4)</td>
<td>335</td>
<td>1x10^{-13}</td>
</tr>
<tr>
<td>Glutamine synthetase type I (EC 6.3.1.2)</td>
<td>129</td>
<td>1x10^{-13}</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (NADP+) (EC 1.4.1.4)</td>
<td>75</td>
<td>1x10^{-13}</td>
</tr>
<tr>
<td>Glutamate dehydrogenase (EC 1.4.1.2)</td>
<td>111</td>
<td>7.6</td>
</tr>
</tbody>
</table>

# aa: length in amino acids
Table 3.6. Transporters identified from BLASTp database from two experimental trials.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Transporters</th>
<th>aa#</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trial-2</td>
<td>Monocarboxylate transporter</td>
<td>116</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>Ammonium transporter</td>
<td>408</td>
<td>0</td>
</tr>
<tr>
<td>Trial-3</td>
<td>Monocarboxylate transporter</td>
<td>118</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>Ammonium transporter</td>
<td>517</td>
<td>0</td>
</tr>
</tbody>
</table>

*aa: length in amino acids
References


CHAPTER 4
FUTURE EXPERIMENTS

4.1 Introduction

Feedlot beef and dairy cattle are fed diets with high levels of grain to increase their production. Due to the fermentative action of specialized microbial groups that digest starch, there is a rapid accumulation of SCFAs from feeding high grain diets, causing a decrease in ruminal pH. Extended periods of low ruminal pH is a condition referred to as acidosis, and this condition can impact animal health and profitability (Kennelly, Robinson, & Khorasani, 1999) (T. Nagaraja, Galyean, & Cole, 1998). Therefore, a great deal of effort has been devoted to the identification of ruminal microorganisms that participate in this process (T. J. DeVries, T. Schwaiger, Beauchemin, & PennerB, 2014).

For many years, culture-dependent approaches were the main tools available to study microorganisms involved in acidosis. While they are very effective at determining the metabolic activities of isolates, one of their main limitations is their inability to provide community-level assessments. In addition, it has been estimated that the vast majority of ruminal microorganisms still remain to be cultured as isolates. The development of sequence-based DNA analysis has provided an effective means of determining community-level assessments of microbial habitats, such as taxonomic composition, diversity, richness, and evenness. Omics-based approaches (metagenomics and metatranscriptomics) can in addition assess the functional potential and activity of microbial communities during acidosis. While these techniques have provided great insights, many microbial mechanisms involved in acidosis remain unknown.
4.2 Research Findings and Future Experiments

Our research resulted in the enrichment of three species-level OTUs from bovine rumen fluid cultures supplemented with starch, and seven species-level OTUs were enriched from bovine rumen fluid cultures supplemented with lactate. Interestingly, one OTU enriched with starch (SD_Bt-01020) and one OTU enriched with lactate (LacT1_D7_3-00877) corresponded to known starch and lactate utilizers, *Ruminobacter amylophilus* and *Coprococcus catus*, respectively. This suggests that the other OTUs enriched in their respective trials were also involved in starch or lactate metabolism.

Since 16S rRNA gene-based analysis was too limited in scope to provide accurate insights regarding the function of candidate starch or lactate-utilizing OTUs, we explored their metabolic potential using metagenomics. Our analysis on specific samples showing enrichment of OTUs in the presence of starch or lactate resulted in the identification of genes predicted to be involved in utilizing the substrate that was supplemented to treatment cultures. Additional functions, such as glycogen synthesis and mobilization, were also uncovered. While metagenomics does not reveal whether predicted genes are active, future studies using metatranscriptomics, for instance, could be conducted to determine whether mRNAs for enzymes of interest are produced, which would indicate that these genes are active. While a major limitation of metagenomics and metatranscriptomics is their dependence on comprehensive databases, i.e. they can only accurately match coding sequences to a known function if enzymes or proteins have been previously reported for this function, our search for candidate enzymes and other proteins involved in carbohydrate metabolism did not appear to be affected. (Aguiar-Pulido et al., 2016).
In my opinion, the most interesting aspect about the *in-vitro* batch culture experimental design is the potential to expand on the project. One way would be to supplement rumen fluid cultures with starch or lactate as continuous cultures, anticipating to see higher relative abundance of candidate starch and lactate utilizers (OTUs), as monitored using 16S rRNA analysis, if a substrate is provided continuously instead of provided as a single dose. For example, I would use seven bio-reactors, with three bio-reactors supplemented with starch, three bio-reactors supplemented with lactate, and one bio-reactor as a non-supplemented control. Starch or lactate substrates would be added at regular intervals to the rumen fluid cultures for the entire trial duration (14-days). As described in previous chapters of this dissertation, I would follow the same methodology to analyze collected samples. Based on the 16S rRNA gene profile, I could determine whether or not the selected candidate starch and lactate utilizers (OTUs) showed higher enrichments. If a difference in the enrichment was observed, my next approach would be to select samples with the highest enrichment for metagenomics analysis. This would identify genes that are important for the selected starch and lactate utilizers (OTUs) to thrive with either starch or lactate as substrates. Since metagenomics analysis is not able to determine whether genes are active or not, I would employ a metatranscriptomics analysis on samples with high enrichment of an OTU to determine which genes are highly expressed. Other strategies to improve levels of enrichment would be to either maintain rumen fluid cultures for longer periods in bioreactors or to perform serial passages of substrate-enriched cultures. For the latter approach, each passage would involve starting a fresh culture using 10% of an enriched culture that would be transferred to sterilized rumen fluid supplemented with the appropriate substrate for enrichment.
Increasing enrichment levels would result in the assembly of longer contigs from metagenomics sequence data, thus providing more accurate predictions of metabolic potential by increasing the number of available annotated genes.

Another possible line of experiments to further pursue this research could be to perform species cultivation. Since cultivation of ruminal bacteria tends to be a challenge, two strategies would be used. First, culture medium would consist of fresh autoclaved rumen fluid mixed with substrate, either starch or lactate, essentially replicating the starting conditions to the bioreactor experiments described in this dissertation. Secondly, the starting pool or source of starch or lactate utilizing bacteria would be from enriched rumen fluid cultures as described above.

4.3 Research impact

Feedlot beef and dairy cattle are fed with grain diets in order to attain greater weight gain and for increased milk production, respectively. Grain overload results in acidosis, which can have a severe impact on the profitability of the beef and dairy industries (Hernández, Benedito, Abuelo, & Castillo, 2014). Starch is the major energy source in grain diets, and its rapid fermentation results in excessive production and accumulation of lactate in the rumen. Various studies have focused on the changes in bacterial diversity, richness, and evenness rather than identifying the key bacterial species, such as amylolytic and lactate fermenting bacteria during lactic acidosis (T. G. Nagaraja & Titgemeyer, 2007) (McCann et al., 2016) (Plaizier et al., 2017).

The research presented in this dissertation was mainly focused on identifying candidate starch and lactate metabolizing bacteria that could grow in batch cultures
consisting of rumen fluid supplemented with starch and lactate. Metagenomics analyses were conducted to determine the functional potential of candidate starch and lactate-enriched bacterial species. Identifying and determining the functional potential of uncharacterized bacteria may give insights into their respective function in the rumen and on how to culture isolates for these species.

One potential long-term impact of this research project may be to provide candidates for the design of probiotics. As described in this dissertation, this research resulted in the identification of only a limited number of candidate starch and lactate utilizers amongst a great deal more that still remain to be identified. However, continuation of this research could lead to the identification of more uncharacterized starch and lactate utilizers (OTUs). In combination with their reconstructed genomes, the uncharacterized species (OTUs) could lead to novel strategies for developing probiotics that are more effective in treating acidosis.
References


