A Comparative Analysis Of The Moose Rumen Microbiota And The Pursuit Of Improving Fibrolytic Systems.

Suzanne Ishaq Pellegrini

University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/graddis

Part of the Animal Sciences Commons, Microbiology Commons, and the Molecular Biology Commons

Recommended Citation


https://scholarworks.uvm.edu/graddis/365

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
A COMPARATIVE ANALYSIS OF THE MOOSE RUMEN MICROBIOTA AND THE PURSUIT OF IMPROVING FIBROLYTIC SYSTEMS.

A Dissertation Presented

by

Suzanne Ishaq Pellegrini

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
Specializing in Animal, Nutrition and Food Science

May, 2015

Defense Date: March 19, 2015
Dissertation Examination Committee:

André-Denis G. Wright, Ph.D., Advisor
Indra N. Sarkar, Ph.D., MLIS, Chairperson
John W. Barlow, Ph.D., D.V.M.
Douglas I. Johnson, Ph.D.
Stephanie D. McKay, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College
ABSTRACT

The goal of the work presented herein was to further our understanding of the rumen microbiota and microbiome of wild moose, and to use that understanding to improve other processes. The moose has adapted to eating a diet of woody browse, which is very high in fiber, but low in digestibility due to the complexity of the plant polysaccharides, and the presence of tannins, lignin, and other plant-secondary compounds. Therefore, it was hypothesized that the moose would host novel microorganisms that would be capable of a wide variety of enzymatic functions, such as improved fiber breakdown, metabolism of digestibility-reducing or toxic plant compounds, or production of functional metabolites, such as volatile fatty acids, biogenic amines, etc.

The first aim, naturally, was to identify the microorganisms present in the rumen of moose, in this case, the bacteria, archaea, and protozoa. This was done using a variety of high-throughput techniques focusing on the SSU rRNA gene (see CHAPTERS 2-5). The second aim was to culture bacteria from the rumen of the moose in order to study their biochemical capabilities (see CHAPTERS 6-7). The final aim was to apply those cultured bacterial isolates to improve other systems. Specifically, bacteria from the rumen of the moose was introduced to young lambs in order to colonize the digestive tract, speed the pace of rumen development, and improve dietary efficiency (see CHAPTER 8).
CITATIONS

Material from this dissertation has been published in the following form:


Material from this dissertation has been submitted for publication to the International Journal of Systemic and Evolutionary Microbiology on December 22, 2014 in the following form:


Material from this dissertation has been submitted for publication to BMC Microbiology on February 23, 2015 in the following form:

Material from this dissertation has been submitted for publication to Applied and Environmental Microbiology on February 11, 2015 in the following form:

ACKNOWLEDGEMENTS

First and foremost, I want to thank my advisor and mentor, Dr. André-Denis Wright, without whom this doctoral project would not have happened. Thanks for supporting me professionally and personally, and especially for endlessly correcting my poor spelling, putting up with my dry sense of humor, and helping me to mature into a professional. Thanks to my other committee members, Dr. John Barlow, Dr. Doug Johnson, Dr. Stephanie McKay, and Dr. Neil Sarkar for their interest in my work, their constructive criticism, helpful suggestions, and for reading every page of this tome. The entire Animal Science Department has my gratitude for too many reasons to list here, but I would like to thank Jane O’Neil and Helen Maciejewski for all of the paperwork they’ve done and phone calls they’ve made on my behalf. I’d like to thank the Wright lab members: the undergraduates who were eager to learn laboratory skills and help with my project, Dr. Benoit St-Pierre and Rachel P. Smith, M.S. for their assistance in technical problems and discussions on analysis, Laura Cersosimo for being my sounding-block and partner in science, and Christina Kim for helping me collect rumen samples from sheep we had to chase down in a field in the hot July sun because literally no one else was available. I’d like to thank my collaborators and those individuals who assisted me in sample collection, as moose parts are not as easy to come by as one would think.
On a more personal note, I’d like to thank my family for their heroic attempts to understand my work and their unfailing support for my nine grueling years of higher education. I especially need to thank my parents, Jill and Tony Pellegrini, for instilling in me a love of reading, organization, and learning, as those have motivated me and proved to be essential while writing this. I’d like to thank my friends, especially Laura Cersosimo, Rachel Smith, Amanda Ochoa, Aimee Benjamin, Mital Pandya, and Mike Haselton, for giving me something to do outside of work, and for always being willing to inevitably talk about work. I want to thank JP Ishaq for his support for so many years, even when I didn’t make it easy, and for putting up with me smelling like a wide variety of animals after work. I’d like to thank everyone who lent a proverbial shoulder to cry on during this last year, and especially for using those shoulders to move all my stuff multiple times up and down stairs. I want to thank Lee Warren for his support in the last and most stressful year of my Ph.D., especially when the only thing to do was to hand me chocolate and hope for the best, and for being a willing participant in my adventures. Finally, I would like to thank the Windows 7 version of Word for its help formatting this beast, and for not being the Windows 8 version of Word.
# TABLE OF CONTENTS

CITATIONS .................................................................................................................. ii
ACKNOWLEDGEMENTS ............................................................................................... iv
LIST OF FIGURES ......................................................................................................... xiii
LIST OF TABLES .............................................................................................................. xv

CHAPTER 1   COMPREHENSIVE LITERATURE REVIEW ........................................... 1
  1.1. Moose .................................................................................................................. 1
    1.1.1 Ecology and anatomy .................................................................................... 1
    1.1.2 Diet and Nutrition ....................................................................................... 3
  1.2. Microbial phylogeny and the small-subunit rRNA genes .................................... 5
  1.3. An overview of the rumen microbiome and its role in digestion ....................... 8
    1.3.1 Bacteria ...................................................................................................... 8
    1.3.2 Archaea and methanogenesis ...................................................................... 11
    1.3.3 Protozoa .................................................................................................... 14
    1.3.4 Fungi ......................................................................................................... 16
    1.3.5 Dietary components and volatile fatty acids ............................................... 17
  1.4. Probiotics and manipulation of the rumen environment ..................................... 22
  1.5. Summary .......................................................................................................... 23
  1.6. References ....................................................................................................... 24
  1.7. Figures ............................................................................................................. 42

CHAPTER 2   INSIGHT INTO THE BACTERIAL GUT MICROBIOME OF
THE NORTH AMERICAN MOOSE (ALCES ALCES) ............................................ 43
  2.1. Abstract ............................................................................................................ 44
  2.2. Background ..................................................................................................... 45
2.3. Results .............................................................................................................................................. 47
  2.3.1 Quantitative Real-Time PCR ........................................................................................................ 47
  2.3.2 PhyloChip Array ................................................................................................................................ 47
  2.3.3 UniFrac analysis ................................................................................................................................. 50
2.4. Discussion ........................................................................................................................................... 51
2.5. Materials and Methods ....................................................................................................................... 58
  2.5.1 Sample Collection ............................................................................................................................. 58
  2.5.2 DNA extraction .................................................................................................................................. 59
  2.5.3 Quantitative Real-Time PCR ........................................................................................................... 59
  2.5.4 PhyloChip ......................................................................................................................................... 60
  2.5.5 Analysis ........................................................................................................................................... 60
2.6. Competing Interests ............................................................................................................................. 62
2.7. Authors’ Contributions ......................................................................................................................... 62
2.8. Acknowledgements ............................................................................................................................. 62
2.9. References ............................................................................................................................................ 63
2.10. Figures ................................................................................................................................................ 67
2.11. Tables ................................................................................................................................................ 73
2.12. Supplemental Material ....................................................................................................................... 75

CHAPTER 3  HIGH-THROUGHPUT DNA SEQUENCING OF THE RUMINAL BACTERIA FROM MOOSE (ALCES ALCES) IN VERMONT, ALASKA, AND NORWAY.  81
  3.1. Abstract .............................................................................................................................................. 82
  3.2. Background ....................................................................................................................................... 83
  3.3. Methods ............................................................................................................................................ 85
3.3.1 Sample Collection ................................................................. 85
3.3.2 DNA extraction and quantification ........................................... 87
3.3.3 Amplicon library preparation .................................................... 87
3.3.4 Sequencing analysis ............................................................... 89
3.4. Results ............................................................................... 91
3.4.1 Classification of taxa by sample location .................................. 91
3.4.2 Statistical analysis of OTUs and clustering ................................. 93
3.5. Discussion .................................................................... 96
3.5.1 Bacteroidetes:Firmicutes ......................................................... 96
3.5.2 Temperature, region, or life stage? ......................................... 99
3.6. Conclusion ..................................................................... 100
3.7. Availability of Supporting Data ................................................. 100
3.8. Competing Interests .............................................................. 100
3.9. Authors’ Contributions ............................................................ 101
3.10. Acknowledgments ................................................................. 101
3.11. References .................................................................. 101
3.12. Figures ......................................................................... 106
3.13. Tables ........................................................................ 109

CHAPTER 4 DESIGN AND VALIDATION OF FOUR NEW PRIMERS FOR
NEXT-GENERATION SEQUENCING TO TARGET THE 18S RRNA GENE OF
GASTROINTESTINAL CILIATE PROTOZOA .................................................. 111

4.1. Abstract ........................................................................ 112
4.2. Introduction .................................................................. 113
4.3. Materials and Methods ....................................................... 115
4.3.1 Sample collection and DNA extraction ................................. 115
4.3.2 Primer design ........................................................................... 115
4.3.3 PCR amplification .................................................................... 116
4.3.4 Sequence analysis ................................................................. 117
4.4. Results ................................................................................... 119
4.4.1 Primer set 1: P-SSU-316F + GIC758R ....................................... 119
4.4.2 Primer set 2: GIC1080F + GIC1578R ........................................ 121
4.4.3 Primer set 3: GIC1184F + GIC1578R ........................................ 122
4.4.4 Comparison of the three primer sets ........................................ 122
4.5. Discussion .............................................................................. 123
4.6. Acknowledgements ................................................................. 127
4.7. References ............................................................................... 127
4.8. Figures ..................................................................................... 132
4.9. Tables ...................................................................................... 135
4.10. Supplemental Material ............................................................. 137

CHAPTER 5  HIGH-THROUGHPUT DNA SEQUENCING OF THE MOOSE RUMEN FROM DIFFERENT GEOGRAPHICAL LOCATION REVEALS A CORE RUMINAL METHANOGENIC ARCHAEAL DIVERSITY AND A DIFFERENTIAL CILIATE PROTOZOAAL DIVERSITY ......................................................... 138
5.1. Abstract .................................................................................. 139
5.2. Introduction ............................................................................... 140
5.3. Methods .................................................................................. 141
5.3.1 Sequence analysis .................................................................. 142
5.3.2 Real-time PCR .......................................................................... 143
5.4. Results ..................................................................................... 144
CHAPTER 6 DESCRIPTION OF STREPTOCOCCUS ALCIS, SP. NOV., AND STREPTOCOCCUS VERMONTENSIS, SP. NOV., AND PROPOSAL OF THREE NEW SUBSPECIES OF STREPTOCOCCUS GALLOYLICUS ISOLATED FROM THE RUMEN OF MOOSE (ALCES ALCES) IN VERMONT

6.1. Summary ................................................................................................................. 164
6.2. Description of Streptococcus alcis sp. nov. .......................................................... 175
6.3. Description of Streptococcus vermontensis sp. nov. .............................................. 176
6.4. Description of Streptococcus galloylicus subsp. mannosilyticus subsp. nov. 177
6.5. Description of Streptococcus galloylicus subsp. melibiosilyticus subsp. nov. 178
6.6. Description of Streptococcus galloylicus subsp. ruminantium subsp. nov. 179
6.7. Conflict of interest ................................................................................................. 179
6.8. Acknowledgements ............................................................................................... 180
6.9. References ............................................................................................................. 180
6.10. Figures ................................................................................................................... 185
6.11. Tables .................................................................................................................... 190
8.4.4 Real-time PCR ................................................................. 228
8.5. Discussion........................................................................... 228
8.6. Acknowledgements............................................................. 230
8.7. References.......................................................................... 231
8.8. Figures................................................................................. 236
8.9. Tables.................................................................................. 244
8.10. Supplemental Material ...................................................... 245

CHAPTER 9  CONCLUSION................................................................. 248
9.1. Summary.............................................................................. 248
9.2. The moose rumen: summarized findings and unanswered questions........ 248
9.3. Fibrolytic probiotics in lambs: beneficial or bust? .......................... 251
9.4. Molecular methodology.......................................................... 255
9.4.1 Sequencing Platform......................................................... 255
9.4.2 Sequencing Target ............................................................. 258
9.4.3 Bioinformatics Programs and Analysis ................................. 260
9.5. References............................................................................ 261
9.6. Figures................................................................................. 267
9.7. Tables.................................................................................. 268

CHAPTER 10  COMPREHENSIVE BIBLIOGRAPHY ......................... 269
LIST OF FIGURES

Figure 1-1 Diagram comparing the immature ruminant to the mature ruminant. ........... 42
Figure 1-2 The frequency of variability in the bacterial 16S rRNA gene [203]. .......... 42
Figure 1-3 Variable regions of the ciliate protozoal 18S rRNA. ............................ 42
Figure 2-1 The OTUs found common in all samples (rumen and colon). ............... 67
Figure 2-2 Breakdown of unclassified families by phylum...................................... 68
Figure 2-3 A comparison of the OTUs exclusive to the rumen or the colon............. 69
Figure 2-4 Jackknife environment clustering in UniFrac, by sample......................... 70
Figure 2-5 Principal component analysis (PCA) scatterplot of the environments using the weighted UniFrac algorithm............................................................... 71
Figure 2-6 Distribution of PhyloChip OTU’s for all 14 samples............................. 72
Figure 3-1 A comparison of sequences per sample, using the Silva bacterial reference taxonomy for classification.......................................................... 106
Figure 3-2 FastUnifrac clustering of all 17 samples. A) weighted, non-normalized, and B) unweighted.......................................................... 107
Figure 3-3 PCoA graphs colored by variable. ......................................................... 108
Figure 4-1 A map of the full-length protozoal 18S rRNA gene, including variable (V1-V9) and rumen ciliate signature regions (SR1-SR4), and showing the respective amplicons of the three primer sets used in the present study. .......... 132
Figure 4-2 Taxonomy and proportion of unique pyrosequences using NCBI (BLAST), by forward primers P-SSU-316F (Sylvester et al., 2004), GIC1080F (present study), and GIC1184F (present study). .................................................. 133
Figure 4-3 Distribution of sequence identity percentages to known sequences based on BLAST by sequencing primer. ...................................................... 134
Figure 5-1 PCoA of moose (A, C, and E) and protozoa (B, D, and F)..................... 157
Figure 5-2 Moose rumen methanogen taxonomy of total sequences, per sample from Alaska, Norway and Vermont .................................................... 158
Figure 5-3 Moose rumen protozoal taxonomy of total sequences, per sample from Norway and Vermont.

Figure 6-1 Neighbor-joining tree comparing isolates to type isolates to lactic acid bacteria.

Figure 6-2 UPGMA trees comparing novel isolates using the housekeeping genes gmk (A), dpr (B), parC (C), pta (D), pyrC (E), recN(F), and rpoD (G).

Figure 7-1 Phylogenetic comparison of 31 isolates which known sequences (NCBI).

Figure 7-2 Growth at various temperatures (A) and pHs (B), as measured by optical density at 600 nm.

Figure 8-1 Group weight (kg) means over time.

Figure 8-2 Group pH means over time.

Figure 8-3 Volatile fatty acid (VFA) and ethanol profile of groups (E=experimental, C=control) for two time points on pasture.

Figure 8-4 Acetate to propionate ratio, with SEM.

Figure 8-5 Principal component analysis of samples by sampling time: May=teal, June=green, July=red, and October=dark blue.

Figure 8-6 Diversity at the phylum level for all sequencing time points.

Figure 8-7 Diversity at the family level for all sequencing time points.

Figure 8-8 Real-time PCR data for methanogens and protozoa.

Figure 8-9 Comparison of methanogen versus protozoal densities for all samples.

Figure 9-1 Mean rumen pH over time, with standard deviation bars.
LIST OF TABLES

Table 2-1 Estimated densities (16S rRNA copy numbers per gram wet weight) of bacteria in the rumen (R) of the moose in October, 2010, Vermont. .................................................. 73

Table 2-2 Total number of taxa found in each sample, before screening for analysis but after background noise was removed and including only OTUs with > 0.92 positive fraction. .................................................................................................................. 74

Table 2-3 Statistics for samples taken from moose shot in October 2010 in Vermont during the moose hunting season................................................................. 74

Table 3-1 Metadata of the 17 samples from Alaska (AK), Norway (NO), and Vermont (VT). .................................................................................................................................................................................. 109

Table 3-2 Statistical measures using a subset of sequences per sample. .................. 110

Table 3-3 The number of shared OTUs across different samples at a 3% genetic distance generated using a shared OTU table in MOTHUR............................................. 110

Table 4-1 Gastrointestinal tract ciliate protozoal primer specifications. .................... 135

Table 4-2 18S rRNA protozoal reference sequences used to determine genetic distance cutoff........................................................................................................... 136

Table 4-3 Comparison of statistical parameters and results for three primer sets....... 137

Table 5-1 Statistical measures per sample for methanogens (met) and protozoa (prot) in Alaska (AK), Norway (NO), and Vermont (VT). ....................................................... 160

Table 5-2 The number of shared OTUs and unique sequences across different samples in Alaska (AK), Norway (NO), and Vermont (VT). ............................................................. 161

Table 5-3 Real-time PCR results for methanogenic archaea and ciliate protozoa in Alaska (AK), Norway (NO), and Vermont (VT).................................................................. 162

Table 6-1 Ability of isolates to produce biogenic amines from selected amino acids, produce lipase, metabolize citrate, produce a ropy or non-ropy exopolysaccharide, and extracellular protein production. ......................................................... 190

Table 6-2 Acid production and ability of isolates to produce an acid byproduct from a variety of carbohydrates. .............................................................................................. 192
Table 6-3 G+C content of isolates and DNA-DNA hybridization to reference strains Streptococcus gallolyticus gallolyticus (ATCC 700065) and S. gallolyticus macedonicus (ATCC BAA-249)............................................................................................................................... 194

Table 7-1 Isolate GenBank ID, closest GenBank match with percent identity, growth on minimal media or on high salinity, and ability to reduce nitrite.................... 212

Table 8-1 Diversity statistics per sample for each of the four sampling time points. .... 244

Table 9-1 A comparison of the Silva bacterial database vs the Ribosomal Database Project (RDP) bacterial reference files in ability to classify 16S rRNA gene bacterial sequences.......................................................... 268
CHAPTER 1  COMPREHENSIVE LITERATURE REVIEW

1.1  Moose

1.1.1  Ecology and anatomy

Moose, *Alces alces*, also known as Eurasian Elk in Europe, are the largest browsing ruminant of the Cervidae (deer) family. They are unique among ruminants, as they do not form herds, but will live individually, with the exceptions being mating season and the first 9 to 10 months of a calf’s life. At maturity they reach upwards of 1.5 to 2 meters at the shoulder, live up to 25 years in the wild, and weigh an average of 360 kg (females) to 450 kg (males). Several subspecies of moose are recognized, for which geographic isolation and adaptation has caused differential characteristics, such as antler shape. For example, moose in Alaska (*A. alces gigas*) and eastern Siberia (*A. alces buturlini*) tend to be larger, with males reaching up to 600 kg, and moose in Scandinavia (*A. alces alces*) have white legs instead of the typical brown.

However, investigations of mitochondrial DNA have revealed conflicting results as to the genetic validity of some subspecies [1–3]. Moose originally migrated to the United States from Asia across the Bering Strait approximately 14,000 to 11,000 years ago. From there, they dispersed across North America and genetic subspecies were eventually established: *A. alces gigas* in Alaska and the Canadian Yukon, *A. a. andersoni* in western Canada and the great lakes region of the US, *A. a. shirasi* from the Rocky Mountains and Colorado to Alberta, Canada, and *A. a. americana* from the great lakes region to the east coast [2]. In a comparison of the four proposed subspecies in North America,
mitochondrial diversity was slightly higher for populations in the center and lower in the peripheral populations along the eastern and western coast (excluding Alaska) [2, 3]. The implication was of a large central population which only relatively recently (as recent as the 1900s) dispersed to peripheral territories, thus the genetic diversity between subspecies was not entirely due to geographic isolation [2, 3].

Moose were traditionally found in most boreal and subarctic areas of the northern hemisphere, but deforestation and over-hunting has reduced their range and, in some areas, their population [4]. They do not thrive in warmer climates, and adults will often lose weight during an unusually hot summer, although only calf weight is adversely affected by overly cold winters [5]. Moose prefer young hardwood forest, deciduous mixed forest, and salt-rich wetland habitats in the summer. Like all ruminants, moose have a specialized digestive system with a four chambered stomach: rumen, reticulum, omasum, and abomasum (Figure 1). The rumen/reticulum fosters a complex consortia of microorganisms (bacteria, archaea, protozoa, fungi, viruses), and the collection of these is known as microbiota. It is these microbiota which ferment plant matter that the animal cannot breakdown on its own [6]. The omasum resorbs water from digesta, and the abomasum secretes pepsin and rennet, and thus functionally resembles the glandular or “true stomach”.

Moose are characterized by having wide mouths and long, flexible tongues that have relatively few taste buds, resulting in browse selection based largely on olfactory cues [7]. Molars and pre-molars are sharper than in many other ruminants, indicating a
specialization towards crushing tougher materials rather than grinding thin grasses [7].
The salivary glands are relatively large for ruminants, increasing in size for the summer,
allowing for excess saliva (specifically serous or enzyme-containing saliva) to pass into
the digestive tract. The saliva of deer also contains tannin-binding proteins [8]. Tannins
are plant-based polyphenols which can bind to proteins in the diet and make them
inaccessible for digestion, thus tannin-binding proteins aid in increasing the digestibility
of the diet.

The rumen is relatively small compared to other browsing ruminants of comparable size,
with a large reticulum capable of filtering larger particles back into circulation for
continued fermentation, a small yet elongated omasum, and an abomasum with unusually
thick mucosa [7]. Openings between stomach chambers are unusually wide and can be
further widened [7]; with the additional saliva this allows faster passage of forage
through the system during summer when food is plentiful. Faster passage of forage
through the digestive system has been shown to reduce methane emissions in domestic
cattle [9].

1.1.2  Diet and Nutrition

Diet selection and a preference for certain plant species can be seen in moose in different
locations, but trends towards certain genera can be seen across all moose. Deciduous or
coniferous leaves, twigs and stems are most often consumed, although moose have been
known to strip bark from ash and maple species. New growth is especially sought out, as
foliage is higher in protein and minerals than grass, and the concentrations of toxic plant
secondary compounds is lower. In western North America, the moose diet is overwhelmingly (75-91%) comprised of willow species (*Salix* spp.), but will also incorporate alder, aspen, and birch [10, 11]. In eastern North America, maple, ash, hemlock, pine, fir, and birch comprised the primary diet of moose [12, 13]. In Scandinavian countries (i.e. Norway, Sweden, Finland), birch and pine tend to dominate the diet, as well as blueberry species [13–17].

Dietary efficiency decreases from summer through autumn, especially with respect to cellulose digestion [18, 19]. Caloric intake decreases from summer into winter as well, not only from reduced forage quality and quantity, but also from decreased production of volatile fatty acids in the rumen, especially propionate and butyrate [18, 19]. Interestingly, moose will voluntarily reduce feed intake in the winter regardless of quality and quantity of feed supplied [20]. Moose commonly lose up to 20% of body weight over winter [21], a cycle which is common to other arctic cervids, such as reindeer.

Moose, especially pregnant cows, show an increased preference for aquatic wetland plant and algae species when they are available during the summer, and specifically for those species with a high salt content, such as green algae, *Spirogyra* sp., and bladderworts, *Utricularis* sp. [22]. An estimated 94-96% of sodium intake for a moose comes from aquatic species eaten during the summer; not only can moose detect salt concentrations as low as 1 mmol (or 100 milli-equivalent/liter), but they appear to have an effective method of sodium retention which reduces excess excretion in urine and feces [22]. In addition
to salt, feeding on aquatic plants provides extra water in the diet which is required for ruminant digestion and peristaltic movement.

1.1.2.1. Modified Diets
There has been some research done on formulated rations for captive moose with variable success [23, 24]. Moose fed on large quantities of grass forages are prone to declining health caused by chronic diarrhea and wasting, which can eventually lead to death [24]. Moose, like all ruminants, are also prone to lactic acidosis or “grain overload” [25]. This is common when an animal switches from a natural cellulose-based diet to a manufactured or otherwise highly-digestible starch-based diet. The sudden change in food type causes a sudden shift in rumen bacterial communities, generally from gram-negative to gram-positive bacteria, and promotes the faster replicating lactic-acid bacterial species which prefer a starch substrate. Excess lactic acid is produced, lowering ruminal pH below pH 5.5, which can kill naturally-occurring populations of microorganisms, such as the rumen protozoa [25]. Not only does this decrease appetite and feed intake, but larger amounts of ruminal acid are transported across the rumen wall and into the bloodstream, leading to a more serious metabolic acidosis.

1.2 Microbial phylogeny and the small-subunit rRNA genes
Metagenomic studies do not usually employ culturing techniques, and many rumen microorganisms are too recalcitrant to culture. Thus, putative identification is made using pairwise gene sequence comparisons to known species. The small subunit of the ribosome (SSU rRNA) provides a good platform for both current molecular methods, but
comparative phylogeny as well as, although the gene itself does not provide any information about the phenotypic functionality of the organism. Overall, the rRNA genes evolve very slowly and, since they are ubiquitous, they can be used for comparison across wide variety of taxa.

Prokaryotes, such as bacteria and archaea, have a 16S rRNA gene which is approximately 1,600 base pairs in length. Eukaryotes, such as protozoa, fungi, plants, animals, etc., have an 18S rRNA gene which is up to 2,300 base pairs in length, depending on the kingdom. In both cases, the S stands for Svedberg Units, or sedimentation rates of the RNA molecule, and is a relative measure of weight and size. Thus, the 18S is larger than the 16S. In both genes, there exist regions which are conserved (identical or near-identical) across taxa, and nine variable regions (V1-V9) [26]. The variable regions are not under functional constraint and are prone to higher evolutionary rates (Figures 2, 3), providing a means for identification and classification through analysis [27–31]. The conserved areas are targets for primers, as a single primer can bind universally (to all or nearly-all) to its target taxa.

In addition to a small subunit, ribosomes also possess a large subunit (LSU rRNA), the 23S rRNA in prokaryotes, and the 28S rRNA in eukaryotes. Eukaryotes have an additional 5.8S subunit which is non-coding, and all small and large units of RNA have associated proteins which aid in structure and function. Taken together, this gives a combined 70S ribosome in prokaryotes, and a combined 80S ribosome rRNA in eukaryotes.
The two main challenges facing high-throughput sequencing are in choosing a target for amplification, and being able to integrate the generated data into an increased understanding of the microbiome of the environment being studied, both of which are discussed further on pages 255-260. High-throughput sequencing can currently sequence thousands to millions of reads which are up to 600 bases in length for amplicons and 1,000 bases in length for genomic DNA (e.g. Roche 454). This has forced studies to choose which variable regions of the rRNA gene to amplify and sequence, and has opened up an arena for debate on which variable region to choose [27].

Additionally, the ability to sequence microorganisms without culturing first has led to the exponential growth of online sequence databases, which have variable amounts of detailed information about the sequence entry. Indeed, many bacterial sequences in GenBank are listed as “unclassified” or “uncultured”, making taxonomic analysis for high-throughput methods difficult. Operational Taxonomic Units (OTUs) are a bioinformatics tool for grouping sequences based on percent identity to known sequences, and in this way sequences can have an assigned level of taxonomy even when the taxonomic resolution is low due to short reads, or when sequences cannot be putatively identified using public databases. Different metagenomic studies also assign OTUs at different taxonomic level (97%, 98%, or 99% for bacteria), which can make analysis across studies difficult.
In recent years, the advances in *de novo* shotgun sequencing has allowed for the large-scale investigation of a variety of microbiomes [32, 33]. While microbiome refers to the collective genetic material or genomes of all the microorganisms in a specific environment, the term is often casually used interchangeably with “microbiota”, or is used to describe only the genetic material of a specific type of microorganism (i.e. “microbiome” instead of “bacterial microbiome”). While the same challenge of sequencing without culturing still applies; in that you can identify which pieces of the puzzle are present but not always how they fit together, shotgun sequencing allows for the entire genome to be sequenced. DNA is enzymatically or physically cut into small pieces which are sequenced, these pieces are assembled into contigs or short sections, which themselves are then assembled to more or less recreate the entire genome. This allows for the identification of putative genes for different enzymes, and detailed comparisons of species across multiple loci instead of just one gene. Naturally, this process comes with its own drawbacks of technical difficulties, extremely high data output, and the logistical challenges of reassembling an entire genome. However, it is an interesting way of identifying form and function in a single method, and is furthering the field of molecular genetics.

### 1.3 An overview of the rumen microbiome and its role in digestion

#### 1.3.1 Bacteria

The most important tool a scientist can possess is curiosity; however, the inventor of the microscope and the “Father of Microbiology” was not a scientist by traditional terms. Antonie van Leeuwenhoek was a draper, a local politician, and a lens-maker in the 1600s.
to early 1700s, and it was using these home-made lenses that he began to observe things on a cellular level. He was the first person to view and describe single-celled organisms, such as bacteria, protozoa, and spermatozoa, which he referred to as “animalcules” or “wee beasties”. To date, there are 30 valid bacterial phyla [34, 35], with several more candidate phyla in use (i.e. OP1, TM7, etc.) [36]; however, only a selection of these are found to colonize the rumen. Some phyla, such as those which contain aquatic or soil bacteria (i.e. Chlorobi or Verrucomicrobia, respectively), are often found in the digestive tract as a result of incidental ingestion, and are thought to be transient members of the rumen. The two main phyla which tend to dominate the gastrointestinal tract (GIT) are Bacteroidetes and Firmicutes.

Currently, the phylum Bacteroidetes contains over 7,000 species, genetically adapted to a wide range of environments, such as soil, water, and the GIT [37]. Bacteria belong to the phylum Bacteroidetes are gram-negative anaerobes or aerobes, and it is generally anaerobic members that belong to the class Bacteroidia (formerly Bacteroides) that are found in all parts of the GIT. As members of the phylum Bacteroidetes possess a wide range of enzymes, especially those which digest carbohydrates or proteins, the species profile for the host GIT is determined by the diet of the host. In the GIT, some of them produce butyrate, which has been implicated in upregulating the GI immune system [38], and can alter toxic or mutagenic compounds [39].

In healthy humans, Bacteroidetes is the dominant phyla [40–42]; however, in obese humans Bacteroidetes bacteria are decreased and Firmicutes is the dominant phylum [43–
This is in contrast to ruminants, which are more likely to have Firmicutes as the dominant phylum [46–53]. However, several studies have identified Bacteroidetes as the dominant phylum in adult ruminants [47, 54], growing ruminants [55], and ruminants transitioning to a high-starch diet [56]. Firmicutes bacteria are gram-positive, often form endospores, and are divided into two major classes: Clostridia and Bacilli. Bacteria belonging to the class Clostridia are strict anaerobes, are also found in soil, and many are characterized as cellulolytic, such as Butyrivibrio spp. [52], Clostridium spp. [57], or Ruminococcus spp. [58]. Within the class Bacilli, the major rumen taxa of interest are cellulolytic Bacillus species [59, 60] or lactic acid bacteria (order Lactobacilliales), such as Lactobacillus spp., Lactococcus spp., Enterococcus spp., and Streptococcus spp. [61].

Other major phyla include Fibrobacteres, which contains cellulolytic bacteria, is common to the GIT [62, 63], and is not well understood as a group as they are difficult to culture. Bacteria belonging to the phylum Proteobacteria are more prevalent in the intestines or colon [46, 64], and many are pathogenic, such as Helicobacter or Campylobacter. However, some species of Proteobacteria have been found to be capable of breaking down plant compounds, such as lignin [65]. Bacteria from the phylum Actinobacteria can also be found in the GIT. Many Actinobacteria species are acetogenic, produce antibiotics or other pharmacologically important compounds [66], or are used to create dairy products, such as Bifidobacterium spp.
1.3.2 *Archaea and methanogenesis*

In 1990, a four page article revolutionized the way we classify microbes [67]. From humble beginnings comes the story of archaea, which could not be correctly classified using the Linnaean system of taxonomy, and not even the prokaryote-eukaryote division could settle the issue. Though studies had already been done on these microorganisms [68], they did not yet have a place on the tree of life. Thus, the Domain level of classification was introduced [67], and the potential for microbial research on archaea was suddenly limitless. The Archaeal domain is divided into two main phyla, Euryarchaeota, representing methanogens and related species, and Crenarchaeota, representing the thermophilic extremophiles, and three presumptive phyla: Korarchaeota, Nanoarchaeota, and Thaumarchaeota [69].

Despite providing a multitude of research opportunities, methanogens were quickly singled out as a potential target for reducing the production of methane from various agricultural and industrial sources. Methane has 25 times more potential for global warming than carbon dioxide, 21 times more if you measure it as the combustion of methane into carbon dioxide [70, 71]. As of 2012 in the US, enteric fermentation from domestic livestock was the second largest source of human-related agricultural methane, accounting for 141 Tg CO\textsubscript{2} Eq/year (equivalent of a million metric tons of carbon dioxide) [70]. Nitrous oxide has a potential of 298 times greater than carbon dioxide, and agricultural soil management creates 204.6 Tg CO\textsubscript{2} Eq emissions, while manure management systems create 17.9 Tg CO\textsubscript{2} Eq/year in the US [70]. Why then have we focused on methane from enteric fermentations?
For domestic livestock, methanogenesis represents a loss of dietary efficiency as compounds such as acetate or hydrogen are sequestered by methanogens instead of being used by the host for production (i.e. live weight gain, milk production, wool production, etc.). Much research has been done on methanogenesis and rumen microbial populations between domestic and wild ruminants, as wild ruminants (bison, elk and deer) are estimated to produce up to 0.37 Tg CO$_2$ Eq/ year [72, 73]. This is a drastically different figure than that for domestic livestock, yet population differences are not the only factor. There are, for example, an estimated 300,000 moose and 25 million white-tailed deer [74] in the US, versus 90 million cattle registered with the USDA [75]. Thus, wild ruminants are presumed to produce less methane based on a presumed higher dietary efficiency and lower production demands.

Ammonia, hydrogen, and carbon dioxide gas are also produced by enteric fermentation, and their partial pressure drives many reactions in the rumen [76]. Methane gas is created when hydrogen, available as free protons, H$_2$ gas or NADH and NADPH cofactors, is used to reduce carbon dioxide. It is thermodynamically favorable, and it prevents the accumulation of hydrogen gas in the digestive tract which can be harmful to both microorganisms and host [76]. In addition to living freely in rumen fluid, attached to particulate matter, or attached GIT epithelia [77], many methanogens can be found attached to the extracellular surface or intracellularly within protozoa as symbiotic way of capturing the hydrogen that the protozoa releases [78–80]. Methanogenic archaea
harness this process to generate energy, although some gamma- and alpha-proteobacteria are methanotrophs, and are capable of using methane as their carbon source.

A few different one- or two-carbon molecules may also be used to provide the carbon dioxide and hydrogen necessary for methanogenesis. Acetate, a volatile fatty acid (VFA) released as a byproduct of enteric fermentation, can be used to form methane along several different enzyme pathways. Interestingly, acetogenesis can also act as a hydrogen sink and will inhibit methanogenesis, though it is not thermodynamically favorable under normal rumen conditions. The hydrogen threshold of a methanogen is up to 100 times lower than for the reductive acetogenesis pathway, thus a large methanogen population will reduce the hydrogen concentration below the threshold of acetogenesis and render it inactive [81, 82].

Formic acid can also be broken down to carbon dioxide and hydrogen, both of which can be used to form methane, by the enzyme hydrogenlyase, known also as hydrogenase or formate hydrogenlyase [68]. Hydrogenlyase can be found in other archaeal species like *Thermococcus*, and has also been studied in the bacterium *Escherichia coli*, which uses the enzyme formate hydrogenlyase (formate dehydrogenase coupled hydrogenase, FDH-MHY) complex to generate electron acceptors under anaerobic conditions, if formate is available [83]. Formate in the rumen is provided through consumption of fruit or honey, dietary supplementation, or bacterial production. Various rumen bacteria, such as *Butyrivibrio, Clostridium, Fibrobacter, Lachnospira, Oxalobacter, Prevotella, Ruminobacter, Ruminococcus, Streptococcus*, or *Succinivibrio* [84] produce formate.
Methanol, a breakdown product of pectin, can also be used for methanogenesis. Regardless of the carbon source and enzyme pathway used, the final step of methanogenesis is catalyzed by the methyl-coenzyme M reductase (mcr) gene and is present in all methanogens [85, 86].

The overwhelming majority of rumen methanogen diversity belongs to the genus *Methanobrevibacter* (*Mbr.*) [87–92]. The two main species identified include *Mbr. smithii* [91, 93, 94], which been shown to improve polysaccharide fermentation by bacteria [95, 96], and *Mbr. ruminantium* [97], which is associated with a lower calorie/high-forage diet [98]. *Methanobrevibacter thaueri* is typically low in ruminants, but was found in reasonably high numbers in impala [99] and moose [100]. *Methanosphaera stadtmanae* strictly uses methanol for methanogenesis, and thus is more prevalent in omnivorous or fruitivorous monogastrics [91, 101, 102], and ruminants with higher pectin diets [99, 100].

1.3.3 Protozoa

After the initial discovery by Antonie van Leeuwenhoek in the 1600s, protozoa were discovered nearly 200 years ago [103–105]. Rumen protozoa represent the largest microbial biomass, although they are not the most abundant, and are important contributors to plant biomass degradation [106–110]. Identification of protozoa in the digestive tract has traditionally been accomplished via microscopic identification [79,
107, 111–117], or other molecular methods [118–121], although a few high-throughput studies have recently been published [100, 122–126].

The rumen protozoa, all of which possess cilia as a means of transport, are divided into two main orders: the Entodiniomorphida, some of which are cellulolytic and have one or two zones or bands of cilia on the anterior end (i.e. *Entodinium, Epidinium*, etc.), and the Vestibuliderida, which are completely covered in cilia, and thus, more motile (i.e. *Dasytricha, Isotricha*). Common rumen protozoa include *Entodinium* spp., *Epidinium* spp., *Eudiplodinium* spp., *Isotricha* spp., *Ophryoscolex* spp., and *Polyplastron multivesiculatum*, [127].

Fibrolytic protozoal species include *Polyplastron multivesiculatum* [110], *Eudiplodinium* spp. [128, 129], *Enoploplastron* spp. [114], and *Diplodinium* sp. [114], to name a few. *Entodinium* spp. are a major source of starch and bacterial digestion in the rumen [130]. While total protozoal counts are higher in concentrate selectors [131], *Entodinium* populations are decreased on a high-concentrate diet versus a roughage diet [131], possible due to a shift in the bacteria populations which the *Entodinium* prey upon. Protozoa are sensitive to changes in pH [132], and factors such as diet [87, 131, 133] and weaning strategy [134] will have an effect on the diversity and quantity of rumen protozoa. Likewise, changes in the protozoal population can alter the density of rumen methanogens which use protozoa as a source for hydrogen in methanogenesis. *Polyplastron, Eudiplodinium, and Entodinium* species have been shown to closely
associate with some methanogens, such as *Methanosphaera stadtmanae* and *Mbr. ruminantium* [135–137].

1.3.4 Fungi

Despite contributing to the breakdown of plant material in the rumen [138–141], relatively little work has been done on identifying rumen fungi or understanding their interactions with other rumen microorganisms. Rumen fungi exist in a plant-associated vegetative state or a free-living zoospore state, which was previously thought to be protozoa [138]. Prior to work by C.G. Orpin in the mid-1970s [142, 143], fungi had not been confirmed to colonize in anaerobic environments.

Currently, all identified anaerobic fungi are classified within the order Neocallimastigales (phylum Neocallimastigomycota) [144]. Genera commonly found in the rumen include *Anaeromyces, Caecomyces, Cyllamyces, Neocallimastix, Orpinomyces, Piromyces*, and *Trichoderma*, all of which are fibrolytic [139, 145]. To date, fungi have been identified in a variety of African ruminants [146], domestic ruminants [140, 146, 147], and herbivorous reptiles [146, 148]. Fungi have also been investigated for use as a probiotic for ruminants [149].

In vitro, co-culturing rumen fungi and the archael *Methanobrevibacter smithii* increased xylanase activity and promoted acetate formation over lactate [96]. Likewise, some strains of lactate-utilizing bacteria stimulated the cellulolytic capabilities of fungi in co-culture [150], while cellulolytic bacteria inhibited it [151].
Typically, fungi are identified using the internal transcribed spacer (ITS) region of non-functional DNA between regions coding for ribosomal subunits. Other potential targets for identification, such as the rRNA genes, nuclear housekeeping genes, or mitochondrial genes, are complicated by poor-quality PCR amplification or sequencing, products which are too large for most current next-generation sequencing techniques, or by low gene variability leading to insufficient resolution for identification [152]. In the case of the phylum Neocallimastigomycota, which encompasses many genera of rumen fungi, mitochondria are entirely lacking.

1.3.5 Dietary components and volatile fatty acids

Dietary fiber is the naturally occurring component of plants which is indigestible to animals that lack the proper digestive enzymes. Fiber is a general term which encompasses pectin, gums, mucilage, cellulose, hemicellulose, and lignin. Both pectin and gums are water-soluble, but those fibers that are found in plant cells walls, such as cellulose, hemicellulose, and lignin, are water-insoluble. Dietary fiber not only provides bulk to the diet and aids in water-retention in the intestines, but in the case of ruminants, plant fibers act as substrates to the fibrolytic microorganisms in the GIT, and a certain percentage of fiber is required in the diet to maintain a healthy, functioning rumen.

Fibrolytic microorganisms use various polysaccharide-digesting enzymes, like cellulases, hemicellulase, glucosidehydrolases, xylanase, etc., to break down plant biomass. Cellulose is a linear polysaccharide of 15-10,000 glucose molecules connected by β-1,4
glycosidic linkages, and provides the structural support in cell walls, which allows for plant growth. Thus, is in the most abundant organic polymer on Earth. As the plant matures and grows in size, the structural components and cellulose content of the plant are increased, leading to a decrease in the digestibility of the plant and lowering its nutritional content. Cellulose can exist as a crystalline structure, as in cotton fibers, or as a manufactured derivative, such as carboxymethylcellulose, which is more water-soluble. Acidic treatment or very high temperatures can make cellulose amorphous and soluble in water.

The crystalline structure can vary depending on the organism which created it. For example, plants create cellulose 1β, which is often mixed with hemicellulososes and lignin to decrease hydrophobia, while bacteria and algae create cellulose 1α, which is found in longer strands and has a higher tensile strength. The cellulase enzyme, often also known as β-1,4 endoglucanase, is actually a group of enzymes. Due to the differing structure of cellulose, all endoglucanase enzymes breakdown cellulose using slightly different actions. In the rumen, bacteria, protozoa, and fungi all produces cellulases [58–60, 106, 110, 139, 145, 153].

Hemicelluloses, also called arabinoxylans, are polysaccharides that make up the plant wall, but are less complex (500-3,000 glucose units) than cellulose, and are found in a branched, amorphous structure. Hemicelluloses include xylan, arabinoxylan, glucuronoxylan, glucomannan, and xyloglucan, and as a group are much more abundant
than cellulose. Hemicellulase enzymes are, likewise, as varied as their substrates and are also produced by rumen bacteria, protozoa and fungi [141, 145, 154–156].

Starch is another plant polysaccharide comprised of glucose units; however, it is more easily digested as both microorganisms and animals possess the required enzymes. Glucose units bound together with $\alpha$-1,4 glycosidic bonds form the helical sugar amylose, which is more resistant to digestion due to its shape. Amylopectin, which is also a polymer of glucose units with $\alpha$-1,4 glycosidic bonds, has branching which occurs at the evenly spaced $\alpha$-1,4 glycosidic bonds. Amylose and amylopectin are the two components of starches, and while amylopectin will make up at least 70% of the starch, there will be different proportions of each depending on the type of starch. Amylose, an enzyme commonly found in saliva and intestines, is able to hydrolyze amylopectin. New growth plants tend to have higher concentrations of starch, which declines as the season progresses and the energy from stored starch is used to facilitate plant growth.

Lignin is an important component of plant cell walls, as it is covalently bonded to hemicellulose, and fills the space in cell walls to provide flexibility and mechanical strength to the plant. Lignin is hydrophobic, allows for water transport through the plants vascular system without allowing it to move osmotically across every cell wall. Although it provides a great deal of carbon, it is of no nutritional value to the ruminant. Rumen bacteria and fungi, some of which are able to breakdown lignin, are able to increase the dietary efficiency of the host and the digestibility of the forage [65, 141, 149, 157].
The products of these plant polysaccharide-targeting enzymes are smaller molecules of cellobiose (glucose dimers) or glucose monomers, which are then fermented by the gut microorganisms. As any free glucose in the rumen is immediately fermented by microorganisms, the host relies on other products for energy. These microbial by-products are primarily short-chain or volatile fatty acids (VFAs) like propionic, acetic, and butyric acids, and to a lesser extent, isobutyrate, valerate, isovalerate, 2-methylbutyric, hexanoic, and heptanoic acids, which are absorbed across the rumen wall and can be used as energy by the host [158–160]. For the most part, these VFAs can be interconverted, and will work towards equilibrium even as new VFAs are introduced into the rumen.

Fiber digestion in the rumen increases the production of the VFA acetate, which is the major VFA found in the rumen and accounts for the increasing proportion of acetate production in the rumen on a higher-forage winter diet [6, 18, 19, 161]. Acetic acid is used for adenosine triphosphate (ATP) synthesis, body fat synthesis, milk synthesis, producing acetyl-CoA for lipogenesis, increasing blood flow to the colon, and other associated health benefits [162].

Sugar, starch, and pectin digestion in the rumen favors the pathway to create the VFA propionate, which accounts for the higher concentrations of propionate production in summer as compared to autumn/winter [6, 18, 19]. Propionate is a biologically important metabolite that is used for gluconeogenesis in the liver, for energy in milk synthesis, it
can induce satiety, and has a variety of other health benefits [163, 164]. The production of propionate reduces the amount of free hydrogen in the rumen, and combined with a drop in the methanogen-commensal protozoal populations associated with high starch diets, this leads to a decrease in methanogenesis and total methanogen populations [133, 165]. Thus, a lower acetate to propionate ratio is indicative of increased dietary efficiency.

Butyrate is an important VFA as it provides energy for the rumen wall itself, as well as intestinal epithelia. It is also used in fat or milk synthesis, and has been shown to reduce inflammation and carcinogenesis in the colon [38, 162, 166, 167]. Butyrate production is increased with a higher fiber diet, and decreased with diets higher in gums [167].

Lactic acid, or lactate, can be used in lieu of glucose as a carbon source by lactose-fermenting bacteria and yeast [168, 169], and is an intermediate step in the production of other VFAs, such as propionate [168]. However, when a rapidly-fermentable carbohydrate, such as some types of starch, is ingested by the ruminant in large quantities, it can lead to a rapid accumulation of lactic acid, which reduces the pH to a point where microorganisms are killed and normal rumen fermentation processes are inhibited [170–173]. Fibrolytic bacteria and protozoa require an environmental pH of 6.3 to 7.0, amylolytic bacteria prefer pH 5.5 to 6.5, and fungi require 6.0 to 7.0.
1.4 Probiotics and manipulation of the rumen environment

Probiotics are live microbial cultures which are administered to promote healthy digestive function and normal microbiota. Probiotics are typically monocultures or comprised of a few species of bacteria [174–179], although yeast [180], and even fungi [149] have been used as probiotics in humans and animals. Transfaunation is the process of transferring rumen contents from one host to another, in contrast to fecal transfer, which uses material from the intestines or feces as an inoculant and is more common in monogastrics. In either case, transferring whole contents allows for the transfer of bacteria, methanogens, protozoa, fungi, and viruses. Prebiotics are chemical or feed additives which promote a healthy digestive function by improving and supporting the growth of normal GIT microbiota [163, 181].

Probiotics and transfaunation have long been used to treat GIT dysfunction, especially after surgery [182–185], but they are also popular as preventative methods to improve overall health [175, 186–189]. For production animals, probiotics have been administered to improve meat, milk, or wool output. This has been accomplished by increasing fibrolysis in the rumen and, thus, dietary efficiency [149, 176–178, 187, 190–194], or by reducing energy wasted in methanogenesis [82] and GIT dysfunction [183].

Rumen development in young ruminants is linked to colonization of the rumen. Colonization begins during birth and progresses through to adulthood [195–198], after which the rumen population may change with diet and health status, but is generally considered to be stable [199, 200]. Host GIT epithelial cells will also adapt to normal
GIT microbiota. Host cells use pattern-recognition receptors to recognize conserved microbial markers, and in response will activate pro- and anti-inflammatory responses as well as upregulate epithelial cell signaling [201, 202]. Thus, alteration of the GIT microbiota is difficult to achieve long-term, especially when the environment is colonized by well-established and well-adapted microorganisms.

1.5 Summary

It was hypothesized that the moose would host novel microorganisms; that these microorganisms would be capable of a wide variety of enzymatic functions; and that these microorganisms could be used to improve other systems. The objectives of this work were to identify the bacteria, archaea, and protozoa present in the rumen of moose; to culture bacteria from the rumen of the moose in order to study their biochemical capabilities; and to apply those cultured bacterial isolates to improve fiber digestion in neonate lambs.
1.6 References


16. Wam HK, Hjeljord O: **Moose summer and winter diets along a large scale gradient of forage availability in southern Norway.** *Eur J Wildl Res* 2010, **56**:745–755.


33. Morrison M, Pope PB, Denman SE, McSweeney CS: Plant biomass degradation by gut microbiomes: more of the same or something new?. *Curr Opin Biotechnol* 2009, **20**:358–363.


77. Leadbetter JR, Breznak JA: Physiological ecology of *Methanobrevibacter cuticularis* sp. nov. and *Methanobrevibacter curvatus* sp. nov., isolated from the hindgut of the termite *Reticulitermes flavipes*. *Appl Env Microbiol* 1996, **62**:3620–3631.


98. Zhou M, Hernandez-Sanabria E, Guan LL: Characterization of variation in rumen methanogenic communities under different dietary and host feed efficiency
conditions, as determined by PCR-denaturing gradient gel electrophoresis analysis. 
*Appl Env Microbiol* 2010, **76**:3776–3786.


102. Facey H V, Northwood KS, Wright A-DG: Molecular diversity of methanogens in fecal samples from captive Sumatran orangutans (*Pongo abelii*). *Amer J Primatol* 2012, **74**:408–413.


185. Lettat A, Benchaar C: Diet-induced alterations in total and metabolically active microbes within the rumen of dairy cows. PLoS One 2013, 8:e60978.


193. Krause DO, Bunch RJ, Conlan LL, Kennedy PM, Smith WJ, Mackie RI, McSweeney CS: Repeated ruminal dosing of Ruminococcus spp. does not result in persistence, but changes in other microbial populations occur that can be measured with quantitative 16S-rRNA-based probes. Microbiol 2001, 147:1719–1729.


1.7 Figures

Figure 1-1 Diagram comparing the immature ruminant to the mature ruminant.

Figure 1-2 The frequency of variability in the bacterial 16S rRNA gene [203].

Figure 1-3 Variable regions of the ciliate protozoal 18S rRNA.
Adapted from Ishaq and Wright, 2014 [122].
CHAPTER 2  

INSIGHT INTO THE BACTERIAL GUT MICROBIOME OF THE NORTH AMERICAN MOOSE (ALCES ALCES).

Suzanne L. Ishaq¹,* and André-Denis G. Wright¹,²

¹ Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, 570 Main St., Burlington, Vermont 05401

² Department of Medicine, University of Vermont, 111 Colchester Ave., Burlington, Vermont, 05401

* Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax) 1-802-656-8196.

Email addresses:
SLI: slpelleg@uvm.edu
ADGW: adwright@uvm.edu

Keywords: colon/gut microbiome/rumen/Vermont/16S rRNA


__________________________________________

43
2.1 Abstract

**Background:** The work presented here provides the first intensive insight into the bacterial populations in the digestive tract of the North American moose (*Alces alces*). Eight free-range moose on natural pasture were sampled, producing eight rumen samples and six colon samples. Second generation (G2) PhyloChips were used to determine the presence of hundreds of operational taxonomic units (OTUs), representing multiple closely related species/strains (>97% identity), found in the rumen and colon of the moose.

**Results:** A total of 789 unique OTUs were used for analysis, which passed the fluorescence and the positive fraction thresholds. There were 73 OTUs, representing 21 bacterial families, which were found exclusively in the rumen samples: Lachnospiraceae, Prevotellaceae and several unclassified families, whereas there were 71 OTUs, representing 22 bacterial families, which were found exclusively in the colon samples: Clostridiaceae, Enterobacteriaceae and several unclassified families. Overall, there were 164 OTUs that were found in 100% of the samples. The Firmicutes were the most dominant bacteria phylum in both the rumen and the colon.

**Conclusions:** Using PhyloTrac and UniFrac computer software, samples clustered into two distinct groups: rumen and colon, confirming that the rumen and colon are distinct environments. There was an apparent correlation of age to cluster, which will be validated by a larger sample size in future studies, but there were no detectable trends based upon gender.
2.2 Background

North American moose, \textit{(Alces alces)}, are the largest browsing ruminant of the deer family Cervidae, and preferably inhabit young hardwood forests, deciduous mixed forests, and salt rich wetland habitats that have an abundance of woody browse and salty aquatic vegetation [1–4]. In northern latitudes, such as Vermont, moose have traditionally done well, although unregulated hunting and deforested habitats caused a severe decline in the Vermont population during the 20\textsuperscript{th} century [5]. It was not until 1993 that moose hunting became regulated again in Vermont and remains strictly controlled by the state. Vermont provides a wide variety of habitats, with one of the most suitable regions being in the northeastern corner of the state. Known as the Northeast Kingdom, the area is rich in bogs and swamps, and is comprised of over 75\% deciduous or mixed forests with growth of various maturities [6]. This area also supports the highest concentration of moose in the state [6] and traditionally has the highest hunter success rates: ranging from 38-70\% from 2006 to 2009 [7, 8], making it an excellent site for sample collection.

Like all ruminants, moose have a specialized digestive system with a four chambered stomach that allows a complex consortium of symbiotic microorganisms to ferment plant matter that the animal cannot breakdown on its own, especially cellulose [9, 10]. During the process of fermentation, hydrogen, ammonia, carbon dioxide, and methane gas are produced [11], as well as volatile fatty acids (VFAs) such as acetate, butyrate, and propionate. These VFAs are released into the rumen where they can be absorbed and used by the ruminant as a source of energy [11–13].
Limited work has previously been done using classical microbiology to identify organisms found in the rumen of moose [14]. One male moose from Alaska was shot in August of 1985, and bacteria which were isolated and characterized consisted of *Streptococcus bovis* (21 strains), *Butyrivibrio fibrisolvens* (9 strains), *Lachnospira multiparus* (7 strains), and *Selenomonas ruminantium* (2 strains) [14].

For the present study, the second generation (G2) PhyloChip (PhyloTech Inc., California) was used to survey rumen and colon samples for the presence and presumptive identification of bacteria. The G2 PhyloChip uses 16S rRNA gene sequences to rapidly type bacteria and methanogens in a mixed microbial sample without the use of cloning or sequencing [15, 16]. The PhyloChip contains approximately 500,000 probes on its surface, representing over 8,400 species of bacteria and roughly 300 species of archaea [17]. There are 11, 25mer, probes that are designed to hybridize to each specific taxon, allowing for specificity in determining taxa present [17]. Depending on what the probes are designed to target, the PhyloChip can be used to differentiate between different serotypes of *Escherichia coli*, or determine the presence of a species regardless of strain. It is already a popular bacterial screening method for air [15], water [18], and soil [19, 20], and has recently gained favor for digestive tract samples [21, 22]. Due to their specificity and sensitivity, DNA microarrays have also been used to categorize diseased and healthy states [22, 23].

The major objectives of the present study were to type the bacteria present in rumen and colonic samples, and to compare these findings with other studies of ruminants and herbivores. Given that moose are large browsing herbivores [3], it was hypothesized that the bacterial populations in the browse-fed wild moose would be more
closely related to bacterial populations found in other browse/forage fed animals. This study reports on the bacteria found in the rumen and colon of the North American moose, as well as how these environments relate to other studies of the gut microbiome in various species.

2.3 Results

2.3.1 Quantitative Real-Time PCR

Mean bacteria cell densities were calculated for each rumen sample using standard curves generated by Bio-Rad’s CFX96 software. Based on a regression line created using the bacterial standards ($R^2 = 0.997$), estimated cell density ranged from $8.46 \times 10^{11}$ to $2.77 \times 10^{12}$ copies of 16S rRNA/g in the rumen (Table 1).

2.3.2 PhyloChip Array

1.1.2.2 Combined rumen and colon

A total of 789 unique OTUs were used for analysis which passed the fluorescence and the positive fraction thresholds. Total numbers for each taxonomic group found are listed for each sample (Table 2), which represent raw data before initial screening. There were 789 total distinct OTUs that were found in all the samples combined; 267 Firmicutes, 225 Proteobacteria, and 72 Bacteroidetes being the major phyla. Not all OTUs were found in every sample, but out the total 789 OTUs there were 164 OTUs, comprising 25 bacterial families, which were found across all 14 samples (Figure 1). The most abundant of these families were unclassified, 25%; Lachnospiraceae, 20%; Clostridiaceae, 16% and Peptostreptococcaceae, 7%. The remaining 21 families
represented less than 4% each of the OTUs found in all 14 samples (Figure 1). The OTUs with unclassified families were then classified by phyla; of the 25% of OTUs with unclassified families, the phyla Firmicutes represented 22%, Proteobacteria and Chloroflexi were 17% each, Bacteroidetes was 15%, and all others represented 5% or less (Figure 2a).

Many of the unclassified sequences were presumptively identified in PhyloTrac, as well as in GenBank, based upon the environment where they were found as most of them are uncultured, thereby providing an interesting, if subjective, means of comparison. Unclassified sequences in the moose were related to a range of environmental sequences including 102 “termite gut clone” OTUs, 20 “rumen clone” OTUs, 20 “forest soil/wetland clone” OTUs, 16 “swine intestine/fecal clone” OTUs, six “human colonic clone” OTUs, six “sludge clone” OTUs, four “penguin dropping clone” OTUs, four “chicken gut clone” OTUs, two “human mouth clone” OTUs and a large number of “soil clone” and “water clone” OTUs from various environments. While many of the forest soil/wetland, soil and water clones may represent transient populations that are picked up from the environment, these data correlate with summer diets of moose in Vermont, namely woody browse in forested areas and aquatic plants found in bogs and marshes.

1.1.2.3. Rumen Samples

The rumen samples contained 575 total OTUs; 192 Firmicutes, 142 Proteobacteria, and 66 Bacteroidetes being the dominant phyla. In the rumen samples, there was a range of 308 to 465 OTUs/sample, and an average of 350 OTUs/sample.
(Table 2). There were 237 OTUs found across all eight rumen samples and, of these, 73 OTUs were exclusive to the rumen, representing 21 families (Figure 3). The OTUs with unclassified families were assigned by phyla (Figure 2b), with the dominant phyla being Bacteroidetes, 27%; Proteobacteria, 19%; and Chloroflexi and NC10 with 11% each. NC10 is a candidate phylum consisting of uncultivated and uncharacterized bacteria that is currently named after the location where the bacteria were sampled, Nullarbor Caves, Australia. All other phyla represented 10% or less of OTUs with unclassified families (Figure 2b). Of the unclassified sequences found exclusively in the rumen, there were 51 termite gut clones, 36 marine, wetland, or waterway sediment clones, 13 fecal or colon clones, 11 rumen clones, nine soil clones, and seven sludge clones.

A previous study on rumen microorganisms in the moose [14] identified *Streptococcus bovis* (21 strains), *Butyrivibrio fibrisolvens* (9 strains), *Lachnospira multiparus* (7 strains), and *Selenomonas ruminantium* (2 strains). The present study found *Streptococcus bovis* strains ATCC 43143 and B315 in every sample except for 1C and 2R. *Butyrivibrio fibrisolvens* and *B. fibrisolvens* strain LP1265 were found in all samples except for 3R, 6R, 2C and 3C, whereas *Butyrivibrio fibrisolvens* strain WV1 was found in 8C only. *Lachnospira multiparus* was not present on the chip. However, all 14 samples did contain *Lachnospira pectinoschiza*, as well as *Selenomonas ruminantium* strains S20 and JCM6582.

1.1.2.4. Colon samples

The colon samples contained a total of 658 OTUs; 248 Firmicutes, 194 Proteobacteria and 46 Bacteroidetes. The colon samples ranged from 307 to 597
OTUs/sample, with an average of 413 OTUs/sample (Table 2). There were 235 OTUs that were found across all six colon samples, and of these, 71 OTUs were exclusive to the colon, representing 22 families (Figure 3). Again, the OTUs with unclassified families were assigned by phyla (Figure 2c), with the dominant phyla being Firmicutes, Proteobacteria and Unclassified, 16% each; Gemmatimonadetes and Chloroflexi, 11% each, and Bacteroidetes, 10%. All other phyla represented 10% or less of OTUs with unclassified families (Figure 2c). Again, many unidentified sequences were listed as uncultured clones by location found. The unidentified sequences found exclusively in the colon were related to 52 “termite gut clone” OTUs, 20 “marine, wetland, or waterway sediment clone” OTUs, 10 “soil clone” OTUs, eight “fecal/colon clone” OTUs, eight “sludge clone” OTUs and five “rumen clone” OTUs.

2.3.3 UniFrac analysis

P-test significance was run using all 14 samples together and 100 permutations, resulting in a corrected p-value of < 0.01, designating that each sample was significantly different from each other. Environment clusters and jackknife values are provided (Figure 4), showing a statistical measurement of the correctness of the tree created. The weighted algorithm accounted for the relative abundance of sequences in a sample, which is typical for environmental samples. UniFrac and PhyloTrac both clustered the rumen and colon samples into two distinct groups: the first node was present 100% of the time in the unweighted and weighted UniFrac clusters. The branching pattern for the rumen group is different between UniFrac algorithm (Figure 4) and between programs (Figure 5). However, the branching pattern for the colon group is identical between PhyloTrac,
and the unweighted and weighted UniFrac outputs. A principal component analysis (PCA) scatterplot (Figure 5) was also created using the weighted algorithm, which grouped the rumen and colon samples separately.

The rumen samples also tentatively clustered by age/weight in the unweighted UniFrac output (Figure 4a), with the youngest/lightest two grouped together (185 kg., 1-yr old; 186.36 kg, 2-yrs old), the two 3-yr old females, grouped together (244.55 and 259.55 kg), and the three oldest/heaviest males (301.36 kg, 4-yrs old; 319.09 kg, 4-yrs old; and 405.45 kg, 8-yrs old) grouped together with a male of unspecified age/weight. The age/weight clusters within the rumen in the weighted UniFrac output (Figure 4b) were not the same as with the unweighted output, nevertheless, some clusters remained (c.f. Figures 5a and 5b).

2.4 Discussion

The major objective of this study was to identify bacteria present in the rumen and colon content samples of the North American moose. This is the first time that the rumen and colon bacterial populations of the moose have been evaluated on a large scale (i.e. PhyloChip), with the last work published in 1986 [14]. While Dehority’s [14] results give the present study an indication of the bacterial population within the rumen of moose, the findings were limited by a sample size of one animal and the constraints of classical microbiology. Anaerobic gut microorganisms are difficult to culture, which continues to present a major obstacle in gut microbial identification. However, genetic analysis, such as microarray and high-throughput sequencing, allow microbes to be studied before they are grown in a pure culture.
One drawback of using the PhyloChip, and indeed with all methods that forego culturing, is the inability to distinguish between live and dead microbes. It also cannot distinguish between colonizing versus transient species, such as the green sulfur bacteria in the phylum Chlorobi or green non-sulfur bacteria of Chloroflexi, both of which are photosynthetic and picked up by the moose during feeding. Careful analysis of the data is required to properly interpret the results. However, even dead and transient bacterial populations can have a profound impact on the resident bacteria as well as the host, whether by releasing harmful components when lysed, such as Lipid A, or providing DNA which may be taken up by live cells in the rumen, as in plasmids that contain genes that confer antibiotic resistance. It is important to take a holistic view to prevent marginalizing potentially important species. Like all methods that rely on PCR amplification, PhyloChip is also subject to PCR bias. This is mediated during sample preparation by running multiple reactions per sample and minimizing the number of cycles.

Rumen samples were consistently clustered separately from the colon samples by PhyloTrac and UniFrac and there were 174 OTUs that were exclusive to either the rumen or the colon; confirming that the rumen and the colon are two distinct environments. Similar findings were reported in a study using fecal samples from sheep [24], as a non-invasive means of modeling the rumen bacteria from captive exotic animals where it is impractical to obtain rumen contents. It was concluded that bacterial concentrations and species in the colon were not reliably predictive of the bacterial concentrations or species in the rumen [24].
The rumen contained an average of $1.66 \times 10^{12}$ copies of 16S rRNA/g ($\pm 7.27 \times 10^{11}$ SEM). This is comparable to other ruminants: $5.17 \times 10^{11}$ cells/g ($\pm 3.49 \times 10^{11}$) for Norwegian reindeer [25], $1.86 \times 10^{11}$ cells/g ($\pm 9.68 \times 10^{10}$) and $5.38 \times 10^{11}$ cells/g ($\pm 2.62 \times 10^{11}$) for Svalbard reindeer [26] in April and October, respectively, and $1.60 \times 10^{11}$ cells/g ($\pm 1.35 \times 10^{11}$) for Canadian dairy cattle [27].

The dominant phylum in the moose rumen was Firmicutes with 192 OTUs, followed by Proteobacteria with 142 OTUs and Bacteroidetes with 66 OTUs. Firmicutes is often the dominant phylum in gut microbiomes, and many of those found in the moose were of the class Clostridia, containing sulfate-reducing bacteria (SRB), which can be pathogenic, endospore forming, and found in soil. Sundset et al. [28] reported that in rumen samples taken from reindeer in Svalbard, the bacteria cultivated were mainly from the class Clostridia. It was noted that *Fibrobacter succinogenes*, *Ruminococcus albus*, and *R. flavefaciens* were not found in the rumen of the reindeer [28], although this may simply be a bias of the cultivation approach. *Fibrobacter* and *Ruminococcus* are both cellulolytic and have previously been found in the rumen of reindeer [25, 29]. However, in the present study, *F. succinogenes* and *R. albus* were not found, despite both species being present on the chip with multiple strains. *Ruminococcus flavefaciens* was detected in several samples, but only a few of its 11 probes matched, making the result insignificant. *Ruminococcus obeum* was detected in the present study.

In a recent paper studying rumen bacteria in dairy cattle, Firmicutes was the dominant phylum in four cattle rumen samples when using full length 16S rRNA clone libraries, but was only dominant in three samples with Proteobacteria being dominant in one sample when using partial 16S rRNA clone libraries or environmental gene tags [30].
Gamma- and alpha-Proteobacteria have been shown to be type I and type II methanotrophs, respectively, meaning they utilize methane as their source of carbon. In the present study, the species Enterobacter cloacae, of the class gamma-Proteobacteria, was found in the moose, and in a non-lactating Holstein cow based on PCR of the 16S rRNA gene to target methanotrophs [31].

In a comparison between the moose rumen data and a study using the PhyloChip and samples from the crop of the wild folivorous bird, the hoatzin [21], similarities arise. Godoy-Vitorino et al. [17] showed that bacteria from the crop of the hoatzin clustered into distinct groups by age: chicks (n=3), juveniles (n=3) and adults (n=3). This correlates with the present study, as the rumen samples clustered by age/weight in the unweighted, and to some extent, in the weighted UniFrac jackknife clustering. As in the moose, some of the differential families found in the crop of the adult hoatzin included Lachnospiraceae, Acidobacteriaceae, Peptostreptococcaceae, Helicobacteraceae and Unclassified (phyla: Proteobacteria, Cyanobacteria, NC10, Chloroflexi, etc.) [17]. The total number of taxonomic groups discovered for hoatzin chicks, juveniles and adults ranged from 37-40 phyla, 47-49 classes, 88-90 orders, 147-152 families, 305-313 subfamilies, and 1351 to 1521 OTUs, an increase over moose, which possibly arises from grouping three samples onto one chip, as was done with the hoatzin samples [21].

In the study by Godoy-Vitorino et al. [21], as well as the current study, OTU cutoff level was predetermined by the PhyloTrac program (i.e. <97%). However, Godoy-Vitorino et al. [17] used a pf=0.90 to determine if an OTU was present, meaning that 90% of the probes for that OTU were positive. When a pf value of 0.90 was applied to the current study, effectively lowering the number of probes that needed to be positive to
be a match for that OTU, the average number of OTUs present rose from 350 to 488 for the rumen and from 413 to 524 for the colon. This suggests that moose either have only a relatively few bacterial species in large quantities, or that there is a wide variety of bacteria found in the moose which are unique and unable to hybridize to the probes found on the G2 PhyloChip. The PhyloChip has recently been shown to overestimate species diversity [32]. The major drawback to using DNA microarray chips is that only known sequences can be used as probes, thus rendering the chips ineffective for discovering and typing new species [33]. The G2 PhyloChip was created in 2006, thus any new taxa that have been identified since then will not be present on the chip, and any re-classification of sequences that are currently on the chip can only be noted by using the most current version of PhyloTrac. These data will be validated and expanded upon using high-throughput DNA sequencing and cultures.

Despite the many similarities between bacteria found in the rumen of the moose to the hoatzin, reindeer and the previous moose study, there are many bacterial families found in the present study which were not mentioned in any of the previous studies. However, many of these bacterial families have been noted in the foregut of the dromedary camel, a pseudo-ruminant with a three chambered stomach. In a recent study by Samsudin et al. [34], the following bacterial families were found in the foregut dromedary camels (n=12) as well as the rumen of the moose in the present study (though not in every rumen sample): Eubacteriaceae, Clostridiaceae, Prevotellaceae, Lachnospiraceae, Rikenellaceae, Flexibacteraceae, Bacteroidaceae, Erysipelotrichaceae, Bacillaceae, Peptococcoceae, and Peptostreptococcaceae. Wild dromedary camels in Australia survive on a high fiber forage diet [34], which is closer to the diet of wild North
American moose. This may explain why the bacterial populations in wild camels appear to be closer to moose than that of wild reindeer, which eat a diet rich in lichens, despite the reindeer and the moose being members of the Cervidae family.

In the rumen, there were 51 sequences found that were listed as being related to termite gut clones, yet many more similarities can be found between the moose and the termite gut, which have compartmentalized guts containing microbes. *Treponema primitia* strain ZAS-1, as well as five other *Treponema* species, were found in the moose rumen in the present study, and 109 *Treponema* phylotypes and species were previously found in the termite gut [35]. *Treponema primitia*, belonging to the phylum Spirochetes, is an acetogenic microorganism capable of degrading mono- and disaccharides such as cellulose or xylan [35]. Bacteroidetes, Chlorobi, Cyanobacteria, Firmicutes and Proteobacteria clones were also discovered in the termite [35], as well 49 phylotypes which represented three new candidate orders in the phylum Fibrobacteres.

To our knowledge, no studies exist using PhyloChip analysis on the fecal samples of herbivores. However, many other colon studies exist, focusing on medically significant pathogens in humans. In a recent study on irritable bowel syndrome, the bacterial families in healthy rats were Rhizobiaceae, Peptococcaceae/Acidaminococcus, Clostridiaceae, Lachnospiraceae, Intrasporangiaceae, Succinivibrionaceae, Alteromonadaceae, Paenibacillaceae and Flavobacteriaceae [36]. Of these, only Peptococcaceae/Acidaminococcus, Clostridiaceae and Lachnospiraceae were found in the moose. In a separate study, fecal samples from cervid species in Norway were tested for colon bacteria that were known pathogens to humans using selective culturing techniques [37]. In that study, *E. coli* O103 was found in 41% of the samples, *E. coli* O26 and O145
were found in small amounts, and *E. coli* O111 and O157 were not found at all [37]. In addition, no cervid fecal samples were positive for *Salmonella*, although one roe deer (*Capreolus capreolus*) sample was positive for *Campylobacter jejuni jejuni* [37]. In the present study, several samples contained *Salmonella*, *E. coli*, or *Campylobacter* species, although no strains of verocytotoxic (e.g. O157:H7) or uropathogenic (e.g. CFT073) *E. coli, Shigella* or *Campylobacter jejuni* were found. However, all of the moose colon samples contained *Citrobacter freundii*, a nitrate reducing bacteria commonly found in the environment, which is known to be an opportunistic pathogen in humans.

The moose colon contained 658 OTUs, of which 248 were Firmicutes and 46 Bacteroidetes. In a 2006 study of the mouse gut microbiome in lean and *ob/ob* obese mice, it was discovered that transfaunation with microorganisms from the obese mouse intestine into the lean mice caused increased weight gain and fat deposition [38]. It is important to note that the bacteria in the obese mice had significantly higher proportions of Firmicutes than Bacteroidetes [38].

The work presented here provides the first insight into the bacterial populations in the digestive tract of the North American moose. While the G2 PhyloChip is an excellent tool for identifying known bacteria, it contains only 300 archaeal sequences, which were not utilized because bacterial-specific primers were used. Furthermore, there is currently no microarray that is designed to identify protozoa or fungi. Next generation (high-throughput) sequencing is needed to validate the bacterial population findings of the present study, as well as identify the protozoal, archaeal and fungal populations present in the moose rumen. The PhyloChip, like all methods that do not rely on culturing, cannot be used to differentiate between transient and colonizing species. It can be assumed that
some species found in the moose are simply passing through the digestive tract, having been picked up from the environment, and are not colonizing the tract. Despite this, these transient bacteria may still have an impact on the dynamics within the rumen, and it is important to take a holistic approach when looking at mixed environmental samples. It is also possible that some of these unclassified bacteria which are presumed transient, such as the soil or water clones, are actually colonizing the moose digestive tract and are simply unique to moose.

2.5 Materials and Methods

2.5.1 Sample Collection

All samples were obtained with permission of licensed hunters through the Vermont Department of Fish and Wildlife. Whole rumen (R) and colon (C) contents were collected from moose shot during the October 2010 moose hunting season in Vermont. Samples were collected by hunters within 2 h, if not sooner, of death and put on ice immediately. Hunters were given a written set of instructions about sample collection, and had been instructed verbally as well, to fill the collection containers with material taken from well inside the rumen and colon, and to seal the container quickly to minimize overexposure to oxygen. Samples were then transferred to the laboratory within 24 h, and stored at -20°C until DNA extraction. A total of eight rumen and six colon samples (Table 3) were collected from eight moose. Twelve of the samples were paired rumen and colon contents from the same animal, and two rumen samples did not have corresponding colon samples. Moose were weighed and aged, by examining the
wear and replacement of the premolars and molars of the lower jar, by Vermont Fish and Wildlife biologists at the mandatory reporting stations.

2.5.2 DNA extraction

Samples were fully thawed, and 0.25 gram aliquots of either rumen content or colonic material, were used for extraction. DNA was extracted from all 14 samples using the repeated bead-beating plus column (RBB+C) method [39], and the QIAamp DNA Stool Mini Kit (QIAGEN, Germantown, Maryland). DNA was quantified using a NanoDrop 2000C Spectrophotometer (ThermoScientific, California), and the purity of the DNA extract was verified using gel electrophoresis to molecular weight. DNA extract was also PCR amplified to test quality and verified using gel electrophoresis to determine correct PCR amplicon length prior to quantitative real-time PCR, or hybridization to the PhyloChip.

2.5.3 Quantitative Real-Time PCR

Real-time PCR was used to calculate bacterial concentrations in each sample, and was performed using a CFX96 thermocycler (Bio-Rad, Hercules, CA), using universal bacterial primers 1114-F (5’-CGGCAACGAGCGCAACCC-3’) and 1275-R (5’-CCATTGTAGCAGCTGTGAGCC-3’) [40]. Each reaction contained 12.5μL of the iQ SYBR Green Supermix kit (Bio-Rad, Hercules, CA): 2.5μl of each primer (40mM), 6.5μL of ddH₂O, and 1μL of the initial DNA extract which was diluted to approximately 10 ng/μL. The external standard for bacteria, as previously described [40], was a mix of Ruminococcus flavefaciens and Fibrobacter succinogenes that were serially diluted over
four logs. The protocol consisted of an initial denaturing at 95°C for 15 min, then 40 cycles of 95°C for 30s, 60°C for 30s, 72° for 1 min. This was followed by a melt curve, with a temperature increase 0.5°C every 10s from 65°C up to 95°C to check for contamination. Data were analyzed using the CFX Manager Software v1.6 (Bio-Rad, Hercules, CA).

2.5.4 PhyloChip

DNA (25-50 ng/μl) was sent to the University of Vermont’s Microarray Core Facility for genotyping using the G2 PhyloChip (PhyloTech Inc., San Francisco, CA). There, the 16S rRNA gene of bacteria was PCR amplified using the universal bacterial primers 27F (5’-AGAGTTTGATCCTGGCTCAG-3’) and 1492R (5’-CTACGGCTACCTTGTTACGA-3’) [41], quantified, fragmented, labeled with biotin, and hybridized according to manufacturer’s proprietary instructions. Each amplified sample was hybridized to its own chip, creating 14 total data sets. The analysis platform used was an Affymetrix 7G scanner, and Gene Chip Operating System (GCOS). Data generated is available online at ARRAYExpress.

2.5.5 Analysis

PhyloChip data were analyzed using the software program PhyloTrac v2.0 (available from www.phylotrac.org). PhyloTrac automatically removed background noise as the average of the two least intense fluorescence signals in each chip quadrant, and used internal standards to create a linear scale to normalize fluorescence intensity with concentration of that sequence in the original sample [17]. The 16S rRNA
sequences on the chip were grouped into Operational Taxonomic Units (OTUs) based on a 97% or greater sequence identity, which was predetermined by the program. For each OTU, there are 11 perfect-match probes, and 11 mismatch probes, which are always analyzed in pairs. For an OTU to be considered a positive match to a probe, the signal intensity must be 1.3X the intensity of the mismatch probe [13]. The positive fraction is a measure of how many perfect-match probes matched out of the total number of probe pairs for that OTU. For this study, a positive fraction of 0.92 was used to determine the presence of an OTU in a sample; for each OTU, 92% of the perfect-match probes were positive. A mean intensity threshold of 100 was used, so that only OTUs with signal intensity greater than that were included in the analysis. All 14 sample files were used in the comparison.

Data were evaluated down to the taxonomic level of family for most analyses since each OTU represented more than one species [32]. A heatmap (Figure 6) showing the presence or absence, and relative intensity of each OTU was created using all 14 samples. Samples were arranged in rows and were clustered on the vertical axis. OTUs were arranged vertically and were clustered on the horizontal axis. Clustering was done using PhyloTrac’s heatmap option with Pearson correlation, a measure of the correlation between two variables, and complete linkage algorithms (farthest neighbor), which clusters based on the maximum distance between two variables.

UniFrac (available from http://bmf2.colorado.edu/unifrac/), an online statistical program, was used to analyze PhyloChip data [42, 43] and to confirm the clustering functions of PhyloTrac. Data were exported from PhyloTrac for analysis using the UniFrac statistical software. P-test significance was run using all 14 environments
together and 100 permutations, to determine whether each sample was significantly different from each other. A p-value of < 0.05 states that the environments were significantly clustered together. Two Jackknife environment clusters were performed using 100 permutations, the weighted and unweighted UniFrac algorithms, and 307 minimum sequences to keep (UniFrac default for the specified conditions). Jackknife counts were provided for each node, representing the number of times out of 100 that a node was present on the tree when the tree was repeatedly rebuilt. A Jackknife percentage of >50% is considered significant. A principal component analysis (PCA) scatterplot was also created using the weighted algorithm, a chart which arranged two potentially related variables into unrelated variables on a graph, revealing underlying variance within the data.

2.6 Competing Interests

The authors declare no competing interests.

2.7 Authors’ Contributions

SI carried out all DNA extraction, PCR, PhyloTrac and Unifrac analysis, and drafted the manuscript. AW conceived of the study and participated in its design, and edited the manuscript. Both authors approved the final manuscript.

2.8 Acknowledgements

The author would like to acknowledge Rachel P. Smith and Dr. Benoit St-Pierre, Department of Animal Science, University of Vermont, for technical advice; the Vermont
Fish and Wildlife Department for their help in sample collection logistics; and Terry Clifford, Archie Foster, Lenny Gerardi, Ralph Loomis, Beth and John Mayer, and Rob Whitcomb for collection of samples.

2.9 References


2.10 Figures

Figure 2-1 The OTUs found common in all samples (rumen and colon). 164 OTUs found common to all samples (n=14). The Unclassified sections are broken down by phyla in Figure 2a.
Figure 2-2 Breakdown of unclassified families by phylum.
(a) OTUs present in all 14 samples. There were 41 OTUs found exclusively in the rumen that were not classified down to the family level. (b) OTUs found exclusively in the rumen. There were 22 OTUs found exclusively in the rumen that were not classified down to the family level. (c) OTUs found exclusively in the colon. There were 19 OTUs found exclusively in the colon that were not classified down to the family level. Several are candidate phyla and are named by where they were discovered: AD3, soil in Virginia and Delaware, USA; OP3 and OP10, now Armatimonadetes, Obsidian Pool hot spring in Yellowstone National Park, USA; NC10, Null Arbor Caves, Australia; TM7, a peat bog in Gifhorn, Germany; WS3, a contaminated aquifer on Wurtsmith Air Force Base in Michigan, USA.
A comparison of the 73 OTUs exclusive in the rumen (n=8) or 71 OTUs exclusive in the colon (n=6), by family. Families with three or more associated OTUs are labeled in the chart; all other families with two or fewer OTUs are labeled via the legend. The Unclassified sections are broken down by phyla in Figure 2b, and 2c, respectively.
Figure 2-4 Jackknife environment clustering in UniFrac, by sample.
(a) An unweighted UniFrac algorithm and (b) a weighted UniFrac algorithm were used, and were not normalized as different evolutionary rates of gene did not need to be accounted for. Jackknife counts for each are provided for each node. The weighted UniFrac algorithm takes into account abundance of sequences, and is better suited to analysis of mixed bacterial samples. Samples are labeled by individual moose (1-8) and sample type (rumen, R or colon, C), and gender, weight and age information is provided in the legend.
Figure 2-5  Principal component analysis (PCA) scatterplot of the environments using the weighted UniFrac algorithm. Samples are labeled by number (1-8), and groups are shown.
Figure 2-6 Distribution of PhyloChip OTU’s for all 14 samples.
Samples (rumen and colon) are arranged in rows and are clustered on the vertical axis (y-axis). OTU’s are arranged vertically and are on the horizontal axis (x-axis). Clustering was done for each using PhyloTrac’s heatmap option with Pearson correlations and complete linkage algorithms.
2.11 Tables

Table 2-1 Estimated densities (16S rRNA copy numbers per gram wet weight) of bacteria in the rumen (R) of the moose in October, 2010, Vermont.

All figures based on calculations using standard curves generated by the Bio-Rad CFX manager program: bacteria (R² = 0.997).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Bacterial copies of 16S rRNA/g (SEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1R</td>
<td>8.46 x 10^{11}</td>
</tr>
<tr>
<td>2R</td>
<td>1.61 x 10^{12}</td>
</tr>
<tr>
<td>3R</td>
<td>2.57 x 10^{12}</td>
</tr>
<tr>
<td>4R</td>
<td>2.02 x 10^{12}</td>
</tr>
<tr>
<td>5R</td>
<td>9.36 x 10^{11}</td>
</tr>
<tr>
<td>6R</td>
<td>1.21 x 10^{12}</td>
</tr>
<tr>
<td>7R</td>
<td>2.77 x 10^{12}</td>
</tr>
<tr>
<td>8R</td>
<td>1.34 x 10^{12}</td>
</tr>
<tr>
<td>Mean (SEM)</td>
<td>1.66 x 10^{12} (7.27 x 10^{11})</td>
</tr>
</tbody>
</table>
Table 2-2 Total number of taxa found in each sample, before screening for analysis but after background noise was removed and including only OTUs with > 0.92 positive fraction. Not all OTUs were found in every sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Phylum</th>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Sub-family</th>
<th>OTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>1R</td>
<td>20</td>
<td>42</td>
<td>59</td>
<td>83</td>
<td>94</td>
<td>367</td>
</tr>
<tr>
<td>2R</td>
<td>21</td>
<td>43</td>
<td>63</td>
<td>90</td>
<td>103</td>
<td>395</td>
</tr>
<tr>
<td>3R</td>
<td>19</td>
<td>38</td>
<td>51</td>
<td>75</td>
<td>83</td>
<td>308</td>
</tr>
<tr>
<td>4R</td>
<td>23</td>
<td>44</td>
<td>58</td>
<td>80</td>
<td>94</td>
<td>374</td>
</tr>
<tr>
<td>5R</td>
<td>23</td>
<td>46</td>
<td>67</td>
<td>97</td>
<td>109</td>
<td>465</td>
</tr>
<tr>
<td>6R</td>
<td>23</td>
<td>43</td>
<td>56</td>
<td>84</td>
<td>97</td>
<td>382</td>
</tr>
<tr>
<td>7R</td>
<td>22</td>
<td>43</td>
<td>57</td>
<td>86</td>
<td>100</td>
<td>379</td>
</tr>
<tr>
<td>8R</td>
<td>23</td>
<td>45</td>
<td>69</td>
<td>98</td>
<td>116</td>
<td>432</td>
</tr>
<tr>
<td>Mean rumen</td>
<td>22</td>
<td>43</td>
<td>60</td>
<td>87</td>
<td>100</td>
<td>350</td>
</tr>
<tr>
<td>1C</td>
<td>16</td>
<td>33</td>
<td>45</td>
<td>63</td>
<td>72</td>
<td>331</td>
</tr>
<tr>
<td>2C</td>
<td>18</td>
<td>36</td>
<td>54</td>
<td>78</td>
<td>90</td>
<td>378</td>
</tr>
<tr>
<td>3C</td>
<td>15</td>
<td>30</td>
<td>40</td>
<td>54</td>
<td>65</td>
<td>307</td>
</tr>
<tr>
<td>6C</td>
<td>17</td>
<td>34</td>
<td>50</td>
<td>72</td>
<td>84</td>
<td>374</td>
</tr>
<tr>
<td>7C</td>
<td>26</td>
<td>49</td>
<td>82</td>
<td>124</td>
<td>146</td>
<td>597</td>
</tr>
<tr>
<td>8C</td>
<td>21</td>
<td>42</td>
<td>66</td>
<td>98</td>
<td>115</td>
<td>488</td>
</tr>
<tr>
<td>Mean colon</td>
<td>19</td>
<td>37</td>
<td>51</td>
<td>82</td>
<td>95</td>
<td>413</td>
</tr>
</tbody>
</table>

Table 2-3 Statistics for samples taken from moose shot in October 2010 in Vermont during the moose hunting season.

<table>
<thead>
<tr>
<th>Moose</th>
<th>Sample location</th>
<th>Sample name</th>
<th>Gender</th>
<th>Weight, dressed carcass (kg)</th>
<th>Approx. age (yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rumen</td>
<td>1R</td>
<td>F</td>
<td>185</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>1C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Rumen</td>
<td>2R</td>
<td>F</td>
<td>244.55</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>2C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Rumen</td>
<td>3R</td>
<td>M</td>
<td>186.36</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>3C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Rumen</td>
<td>4R</td>
<td>M</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>Rumen</td>
<td>5R</td>
<td>M</td>
<td>319.09</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Rumen</td>
<td>6R</td>
<td>F</td>
<td>259.55</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>6C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Rumen</td>
<td>7R</td>
<td>M</td>
<td>301.36</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>7C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Rumen</td>
<td>8R</td>
<td>M</td>
<td>405.45</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Colon</td>
<td>8C</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### 2.12 Supplemental Material

**Supplementary Table 1** Genus/Identifier and GenBank # of sequences in selected families, found in all rumen samples (n=8), sequences are non-exclusive to the rumen.

<table>
<thead>
<tr>
<th>Rumen</th>
<th>Genus or Identifier</th>
<th>GenBank ASCN #</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clostridiaceae</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Clostridium</em></td>
<td>X71852.1, AB093546.1</td>
</tr>
<tr>
<td></td>
<td>cow rumen clone</td>
<td>AY244908.1</td>
</tr>
<tr>
<td></td>
<td>Great Artesian Basin clone</td>
<td>AF407695.1</td>
</tr>
<tr>
<td></td>
<td>Forested wetland clone</td>
<td>AF523923</td>
</tr>
<tr>
<td></td>
<td>Lachnospiraceae</td>
<td>AF550610.1</td>
</tr>
<tr>
<td></td>
<td>Oral clone</td>
<td>AF481208.1</td>
</tr>
<tr>
<td></td>
<td>sludge clone</td>
<td>AF482434.1</td>
</tr>
<tr>
<td></td>
<td>Swine intestine clone</td>
<td>AF371790.1, AF371796.1</td>
</tr>
<tr>
<td></td>
<td>Termite gut clone</td>
<td>AB100478.1, AB100483.1, AB100493.1, AB100479.1, AB100476.1, AB100486.1, AB089030.1, AB089043.1, AB089045.1, AB088977.1, AB089028.1, AB088965.1, AB088951.1, AB089035.1, AB088984.1, AB089032.1</td>
</tr>
<tr>
<td></td>
<td>Lachnospiraceae</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Butyrivibrio</em></td>
<td>U41168.1, X89978.1, AF105403.1, AY178635.1</td>
</tr>
<tr>
<td></td>
<td>chicken clone</td>
<td>AF376201.1, AF376218.1</td>
</tr>
<tr>
<td></td>
<td><em>Clostridium</em></td>
<td>AF067965.1, AY169415.1</td>
</tr>
<tr>
<td></td>
<td><em>Eubacterium</em></td>
<td>AB008552.1</td>
</tr>
<tr>
<td></td>
<td><em>Fusobacterium</em></td>
<td>X85022.1</td>
</tr>
<tr>
<td></td>
<td>Granular sludge clone</td>
<td>AF332711.1, AF332720.1, AF332721.1</td>
</tr>
<tr>
<td></td>
<td><em>Lachnospira</em></td>
<td>AY169414.1</td>
</tr>
<tr>
<td></td>
<td><em>Pseudobutyrivibrio</em></td>
<td>AF202260.1</td>
</tr>
<tr>
<td></td>
<td>Swine intestine clone</td>
<td>AF371584.1, AF371541.1, AF371648.1</td>
</tr>
<tr>
<td></td>
<td>Rumen clone</td>
<td>AB034003.1, AB034059.1</td>
</tr>
<tr>
<td></td>
<td>Termite gut clone</td>
<td>AB100463.1, AB089002.1, AB088983.2, AB088950.1, AB088993.1, AB088990.1, AB088998.1, AB088994.1, B089044.1, AB088968.1, AB088980.1, AB088952.1, AB089040.1, B088991.1, AB089034.1</td>
</tr>
<tr>
<td></td>
<td>Peptostreptococcaceae</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Anaerococcus</em></td>
<td>AF542229.1</td>
</tr>
<tr>
<td></td>
<td><em>Finegoldia</em></td>
<td>AB109771.1</td>
</tr>
<tr>
<td></td>
<td>municipal wastewater treatment plant clone</td>
<td>CR933145.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------------------</td>
<td></td>
</tr>
<tr>
<td>oral clone</td>
<td>AF538856.1, Y134903.1, AY134904.1</td>
<td></td>
</tr>
<tr>
<td><em>Peptostreptococcus</em></td>
<td>AF481225.1</td>
<td></td>
</tr>
<tr>
<td>TCE-dechlorinating clone</td>
<td>Y217429.1</td>
<td></td>
</tr>
<tr>
<td>termite gut clone</td>
<td>AB100461.1, AB062845.1, AB088986.1, AB088954.2, AB088971.1, AB088960.2, AB088970.1</td>
<td></td>
</tr>
<tr>
<td><strong>Prevotellaceae</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cow rumen clone</td>
<td>AY244919.1, AY244931.1, AY244923.1, AY244900.1</td>
<td></td>
</tr>
<tr>
<td>Deep marine sediment clone</td>
<td>AY093462.1</td>
<td></td>
</tr>
<tr>
<td><em>Prevotella</em></td>
<td>AY699286.1, AY134905.1, AF481227.1</td>
<td></td>
</tr>
<tr>
<td>Rumen clone</td>
<td>AB185583.1</td>
<td></td>
</tr>
<tr>
<td>Swine intestine clone</td>
<td>AF371893.1</td>
<td></td>
</tr>
<tr>
<td><strong>Unclassified</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthrobacter</td>
<td>X80744.1</td>
<td></td>
</tr>
<tr>
<td>Bacillus</td>
<td>AF454298.1</td>
<td></td>
</tr>
<tr>
<td>Beta proteobacterium, uncultured</td>
<td>AY082479</td>
<td></td>
</tr>
<tr>
<td>chicken cecum clone, uncultured</td>
<td>AF376430</td>
<td></td>
</tr>
<tr>
<td><em>Chlorobi</em>, uncultured</td>
<td>AY118151.1</td>
<td></td>
</tr>
<tr>
<td>coal effluent wetland clone</td>
<td>AF523883.1, AF523884.1</td>
<td></td>
</tr>
<tr>
<td>cow rumen clone</td>
<td>AY244902.1</td>
<td></td>
</tr>
<tr>
<td><em>Cryptoendolith</em></td>
<td>AY250886.1</td>
<td></td>
</tr>
<tr>
<td>deep marine sediment clone</td>
<td>AY093480.1</td>
<td></td>
</tr>
<tr>
<td>Delaware River estuary clone</td>
<td>AY274839.1</td>
<td></td>
</tr>
<tr>
<td><em>Desulfitomaculum</em></td>
<td>AY084078.1, Y11573.1, Y11574.1</td>
<td></td>
</tr>
<tr>
<td>Elbe river clone</td>
<td>AJ401113.1</td>
<td></td>
</tr>
<tr>
<td>Feedlot manure, uncultured</td>
<td>AF317384.1</td>
<td></td>
</tr>
<tr>
<td><em>Flexistipes</em></td>
<td>AF481216.1</td>
<td></td>
</tr>
<tr>
<td>forest soil clone</td>
<td>AY913277.1</td>
<td></td>
</tr>
<tr>
<td>forested wetland clone</td>
<td>AF523973.1, AF524015.1</td>
<td></td>
</tr>
<tr>
<td>G+C Gram-positive clone</td>
<td>AF465653.1</td>
<td></td>
</tr>
<tr>
<td><em>Holophaga</em></td>
<td>AJ519640</td>
<td></td>
</tr>
<tr>
<td>human mouth clone</td>
<td>AY207065.1, AY134895.1</td>
<td></td>
</tr>
<tr>
<td>hydrothermal vent sediment clone</td>
<td>AF420340.1</td>
<td></td>
</tr>
<tr>
<td>marine sediment above hydrate ridge clone</td>
<td>AJ535231.1</td>
<td></td>
</tr>
<tr>
<td>Mono Lake clone</td>
<td>AF507879.2, AF507860.1, AF507869.2, AF507859.1, AF507862.1, AF507892.1</td>
<td></td>
</tr>
<tr>
<td>Environmental Sample</td>
<td>Accession Numbers</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Ocean soil, uncultured</td>
<td>AY181050.1</td>
<td></td>
</tr>
<tr>
<td><em>Paralvinella</em></td>
<td>AJ441239.1, AJ441217.1</td>
<td></td>
</tr>
<tr>
<td>Penguin droppings sediments clone</td>
<td>AY218551.1</td>
<td></td>
</tr>
<tr>
<td><em>Pleurotus</em></td>
<td>AY838556.1</td>
<td></td>
</tr>
<tr>
<td>Rocky Mountain alpine soil clone</td>
<td>AY192275.1, AY192273.1</td>
<td></td>
</tr>
<tr>
<td>Rumen clone</td>
<td>AB185532.1</td>
<td></td>
</tr>
<tr>
<td>Sludge clone</td>
<td>AB106352.1, AF234699.1, AF368184.1</td>
<td></td>
</tr>
<tr>
<td>Soil clone</td>
<td>AJ390463.1</td>
<td></td>
</tr>
<tr>
<td>Sponge clone</td>
<td>AJ347025.1, AJ347055.1</td>
<td></td>
</tr>
<tr>
<td>Temperate estuarine mud clone</td>
<td>AY216458.1</td>
<td></td>
</tr>
<tr>
<td>Termite gut homogenate clone</td>
<td>AB088930.1, AB089097.1, AB089109.1, AB089122.1, AB089008.1</td>
<td></td>
</tr>
<tr>
<td>Toolik Lake clone</td>
<td>AF534435.1</td>
<td></td>
</tr>
<tr>
<td>Trichloroethene-contaminated site clone</td>
<td>AF529128.1</td>
<td></td>
</tr>
<tr>
<td>Vadose (subterranean water) clone</td>
<td>AY177763.1</td>
<td></td>
</tr>
<tr>
<td>Vent worm enrichment clone</td>
<td>AJ431235.1</td>
<td></td>
</tr>
</tbody>
</table>
**Supplementary Table 2** Genus/Identifier and GenBank # of sequences in selected families, found in all colon samples (n=6), sequences are non-exclusive to the colon.

<table>
<thead>
<tr>
<th>Genus or Identifier</th>
<th>GenBank ASCN #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clostridiales</strong></td>
<td></td>
</tr>
<tr>
<td>Clostridium</td>
<td>AY007244.1, X76750.1, X71852.1, AB093546.1</td>
</tr>
<tr>
<td>Cow rumen clone</td>
<td>AY244908.1</td>
</tr>
<tr>
<td>Clostridales oral clone</td>
<td>AF481208</td>
</tr>
<tr>
<td>Equine intestine clone</td>
<td>AJ408137.1</td>
</tr>
<tr>
<td>Forested wetland clone, uncultured</td>
<td>AF523923</td>
</tr>
<tr>
<td>Granular sludge clone</td>
<td>AY261814.1, AF482434.1</td>
</tr>
<tr>
<td>Great Artesian Basin clone</td>
<td>AF407695.1</td>
</tr>
<tr>
<td><strong>Lachnospiraceae</strong></td>
<td></td>
</tr>
<tr>
<td>Swine intestine clone</td>
<td>AF371796.1, AF371790.1, AF371783.1</td>
</tr>
<tr>
<td>Termite gut clone</td>
<td>AB100493.1, AB100478.1, AB088977.1, AB100475.1, AB100479.1, AB089043.1, AB089028.1, AB100486.1, AB088951.1, AB088965.1, AB089045.1, AB089032.1, AB100469.1, AB100483.1, AB089035.1, AB089033.1, AB088984.1, AB100476.1, AB089030.1</td>
</tr>
<tr>
<td><strong>Enterobacteriaceae</strong></td>
<td></td>
</tr>
<tr>
<td>Alterococcus</td>
<td>AF075271.2</td>
</tr>
<tr>
<td>Citrobacter</td>
<td>AF025365.1</td>
</tr>
<tr>
<td>Coal effluent wetland clone</td>
<td>AF523903.1</td>
</tr>
<tr>
<td>Enterobacter</td>
<td>AJ550468.1</td>
</tr>
<tr>
<td>Erwinia</td>
<td>AF141891.1, AF373202.1</td>
</tr>
<tr>
<td>Escherichia</td>
<td>NC_000913.2</td>
</tr>
<tr>
<td>Klebsiella</td>
<td>X93216.1</td>
</tr>
<tr>
<td>Kluyvera</td>
<td>AJ627202.1</td>
</tr>
<tr>
<td>Opitutus</td>
<td>AY695840.1</td>
</tr>
<tr>
<td>Pantoea</td>
<td>AF130912.1, AF373198.1</td>
</tr>
<tr>
<td>Raoultella</td>
<td>AF181574.1</td>
</tr>
<tr>
<td>Salmonella</td>
<td>U92194.1</td>
</tr>
<tr>
<td>Serratia</td>
<td>AF124036.1</td>
</tr>
<tr>
<td>Soda lake clone, uncultured</td>
<td>AF507000.1</td>
</tr>
<tr>
<td>White-tail deer rumen clone</td>
<td>AF084835.1</td>
</tr>
</tbody>
</table>

Lachnospiraceae
<table>
<thead>
<tr>
<th>Genus</th>
<th>Accession Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Butyrivibrio</em></td>
<td>U41168.1</td>
</tr>
<tr>
<td><em>Clostridium</em> *</td>
<td>AY169415.1</td>
</tr>
<tr>
<td>Chicken clone, uncultured</td>
<td>AF376205.1, AF376218.1, AF376201.1</td>
</tr>
<tr>
<td><em>Fusobacterium</em></td>
<td>X85022.1</td>
</tr>
<tr>
<td>granular sludge clone</td>
<td>AF332721.1, AF332720.1, AF332711.1</td>
</tr>
<tr>
<td>human colonic clone</td>
<td>AJ408972.1, AJ408989.1</td>
</tr>
<tr>
<td><em>Lachnospira</em></td>
<td>AY169414.1</td>
</tr>
<tr>
<td><em>Roseburia</em></td>
<td>AY804149.1</td>
</tr>
<tr>
<td>Rumen clone</td>
<td>AB034059.1, AB034003.1</td>
</tr>
<tr>
<td>Ruminantium</td>
<td>AB008552.1</td>
</tr>
<tr>
<td>Swine intestine clone</td>
<td>AF371584.1, AF371541.1, AF371648.1</td>
</tr>
<tr>
<td>Termite gut homogenate clone</td>
<td>AB088950.1, AB089040.1, AB088950.1, AB088990.1, AB088998.1, AB088993.1, AB100463.1, AB088994.1, AB089002.1, AB089000.1, AB089044.1, AB088952.1, AB089036.1, AB089034.1, AB088983.2, AB088980.1, AB088968.1, AB088991.1</td>
</tr>
<tr>
<td>Peptostreptococcaceae</td>
<td></td>
</tr>
<tr>
<td><em>Anaerococcus</em></td>
<td>AF542229.1</td>
</tr>
<tr>
<td><em>Finegoldia</em></td>
<td>AB109771.1</td>
</tr>
<tr>
<td>municipal wastewater treatment plant clone</td>
<td>CR933145.1</td>
</tr>
<tr>
<td>Oral clone</td>
<td>AY134904.1</td>
</tr>
<tr>
<td><em>Peptostreptococcus</em></td>
<td>AF481225.1</td>
</tr>
<tr>
<td>Swine manure clone</td>
<td>AY167963.1</td>
</tr>
<tr>
<td>TCE-dechlorinating clone</td>
<td>AY217429.1</td>
</tr>
<tr>
<td>Termite gut clone</td>
<td>AB088971.1, AB062845.1, AB088954.2, AB088986.1, AB088970.1</td>
</tr>
<tr>
<td>Unclassified</td>
<td></td>
</tr>
<tr>
<td>Anaerobic bioreactor clone</td>
<td>AJ278169.1</td>
</tr>
<tr>
<td>Antarctic sediment clone</td>
<td>AY250886.1, AY133397.1, AY177804.1</td>
</tr>
<tr>
<td><em>Arthrobacter</em></td>
<td>X80744.1</td>
</tr>
<tr>
<td><em>Bacillus</em></td>
<td>AF454298.1</td>
</tr>
<tr>
<td>Bacteroidetes, uncultured</td>
<td>AF449785.1</td>
</tr>
<tr>
<td>Beta proteobacterium, uncultured</td>
<td>AY082479</td>
</tr>
<tr>
<td>Brackish mud clone, uncultured</td>
<td>AF211303.1</td>
</tr>
<tr>
<td>Chicken cecum, uncultured</td>
<td>AF376430</td>
</tr>
<tr>
<td><em>Chlorobi</em> clone, uncultured</td>
<td>AY118151.1</td>
</tr>
<tr>
<td>cow rumen clone</td>
<td>AY244902.1, AB185532.1</td>
</tr>
<tr>
<td>DCP-dechlorinating clone</td>
<td>AJ306793.1, AJ306749.1</td>
</tr>
<tr>
<td>Delaware River estuary clone</td>
<td>AY274839.1</td>
</tr>
<tr>
<td>Bacterial Name</td>
<td>Accession Numbers</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Desulfotomaculum</td>
<td>Y11574.1, Y11573.1, AY084078.1</td>
</tr>
<tr>
<td>Feedlot manure clone, uncultured</td>
<td>AF317384.1</td>
</tr>
<tr>
<td>Ferribacter</td>
<td>AF282254.1, AF282252.1</td>
</tr>
<tr>
<td>Forest soil clone, uncultured</td>
<td>AY913277.1, AF507760.1</td>
</tr>
<tr>
<td>forested wetland clone</td>
<td>AF523973.1, AF524015.1</td>
</tr>
<tr>
<td>Holophaga</td>
<td>AJ519640</td>
</tr>
<tr>
<td>hot spring clone</td>
<td>AF027097.1</td>
</tr>
<tr>
<td>Human mouth clone, uncultured</td>
<td>AF515500.1, AY207065.1</td>
</tr>
<tr>
<td>hydrothermal sediment clone</td>
<td>AF420340.1, U15103.1</td>
</tr>
<tr>
<td>Mono Lake clone</td>
<td>AF507859.1, AF507869.2, AF507892.1, AF507879.2</td>
</tr>
<tr>
<td>Ocean sediment clone, uncultured</td>
<td>AY181050.1, AY114325.1, AY093473.1</td>
</tr>
<tr>
<td>Oral clone</td>
<td>AY134895.1</td>
</tr>
<tr>
<td>Paralvinella</td>
<td>AJ441239.1</td>
</tr>
<tr>
<td>Penguin droppings sediment clone</td>
<td>AY218551.1</td>
</tr>
<tr>
<td>Plectonema</td>
<td>AF091110.1</td>
</tr>
<tr>
<td>Rocky Mountain alpine soil clone</td>
<td>AY192273.1, AY192275.1</td>
</tr>
<tr>
<td>Salmonella</td>
<td>AF029226.1</td>
</tr>
<tr>
<td>sludge clone</td>
<td>AF234699.1</td>
</tr>
<tr>
<td>soil clone</td>
<td>AJ390463.1</td>
</tr>
<tr>
<td>sponge clone</td>
<td>AJ347055.1</td>
</tr>
<tr>
<td>temperate estuarine mud clone</td>
<td>AY216458.1</td>
</tr>
<tr>
<td>termite gut clone</td>
<td>AB089097.1, AB089109.1, AB089074.1, AB089008.1</td>
</tr>
<tr>
<td>vent worm enrichment clone</td>
<td>AJ431234.1, AJ431235.1</td>
</tr>
</tbody>
</table>

* This entry is included under family Lachnospiraceae in PhyloTrac, and as family Clostridiaceae in GenBank.
CHAPTER 3  HIGH-THROUGHPUT DNA SEQUENCING OF THE RUMINAL BACTERIA FROM MOOSE (*ALCES ALCES*) IN VERMONT, ALASKA, AND NORWAY.

Suzanne L. Ishaq¹,* and André-Denis Wright¹,²

*¹ Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, Burlington, Vermont 05401

² Department of Medicine, University of Vermont, 111 Colchester Ave., Burlington, Vermont, 05401

*Contact: Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax) 1-802-656-8196.

Keywords: 16S rRNA/454 Roche Titanium/Alaska/Norway/rumen/Vermont

3.1 Abstract

In the present study, the rumen bacteria of moose (Alces alces) from three distinct geographic locations were investigated. Moose are large, browsing ruminants in the deer family, which subsist on fibrous, woody browse, and aquatic plants. Subspecies exist which are distinguished by differing body and antler size, and these are somewhat geographically isolated. Seventeen rumen samples were collected from moose in Vermont, Alaska and Norway, and bacterial 16S rRNA genes were sequenced using Roche 454 pyrosequencing with Titanium chemistry. Overall, 109,643 sequences were generated from the 17 individual samples, revealing 33,622 unique sequences. Members of the phylum Bacteroidetes were dominant in samples from Alaska and Norway, but representatives of the phylum Firmicutes were dominant in samples from Vermont. Within the phylum Bacteroidetes, Prevotellaceae was the dominant family in all three sample locations, most of which belonged to the genus Prevotella. Within the phylum Firmicutes, the family Lachnospiraceae was the most prevalent in all three sample locations. The data set supporting the results of this article is available in the Sequence Read Archive (SRA), available through NCBI [study accession number SRP022590]. Samples clustered by geographic location and by weight, and were heterogeneous based on gender, location and weight class (p<0.05). Location was a stronger factor in determining the core microbiome than either age or weight, but gender did not appear to be a strong factor. There were no shared operational taxonomic units across all 17 samples, which indicates that these moose may have been isolated long enough to preclude a core microbiome among moose. Other potential factors discussed include differences in climate, food quality and availability, gender, and life cycle.
3.2 Background

The moose (*Alces alces*), also known as the Eurasian elk in Europe, is a large browsing ruminant in the Cervidae (deer) family native to northern latitudes, especially Canada, the northern United States, Scandinavia and Russia. Their diet typically includes woody browse from a variety of hardwoods and deciduous species (i.e. willow, aspen, ash, maple) [1, 2], but during warmer months will also include salt-rich aquatic vegetation from wetlands and swamps [3]. As a ruminant, they possess a four-chambered stomach (rumen, reticulum, omasum, and abomasum), of which, the rumen and reticulum contain a diverse assemblage of microorganisms that break down their fibrous diet and provide usable nutrients for the host [1, 4].

As many populations of moose are geographically isolated, there are several recognized subspecies based upon color, structure of the antlers, and facial features/structure. The present study investigated samples from three different subspecies of moose in Vermont (VT) (*Alces alces americana*), Alaska (AK) (*Alces alces gigas*), and northern Norway (NO) (*Alces alces alces*). *Alces alces gigas* is recognized as the largest moose subspecies (>600kg for males, >450kg for females), *A. alces americana* is mid-range (>450kg for males, >250kg for females), and *A. alces alces* tends to be smaller (>250kg for males, >200kg for females). *Alces alces alces* also have more distinctive white legs, and have shorter and broader antlers than the other two subspecies. Generally, moose wean at 6 months, reach sexual maturity at just over 2 years old, and are known to live up to 15-20 years in the wild [4].
Only two published studies currently exist which identify the bacteria present in the rumen of the moose. The first used traditional culturing techniques to identify the bacteria in a male moose from Alaska, and identified *Streptococcus bovis*, *Butyrivibrio fibrisolvens*, *Lachnospira multiparus*, and *Selenomonas ruminantium* [5]. The second study used the second generation (G2) PhyloChip to identify hundreds of bacteria from the rumen and colon of eight moose from Vermont, and reported that the phylum Firmicutes was dominant, followed by the phylum Proteobacteria [6].

The present study compares three geographic locations of moose using Roche 454 high-throughput sequencing of the 16S rRNA gene using a 500bp amplicon generated from the universal bacterial primers 27F [7] and 519R [8]. The objectives of this research were to classify the bacteria present in the rumen of moose from Alaska, Vermont and Norway; to compare the samples across geographic location, gender, and weight class to determine possible trends; and to compare samples to published studies on wild and domesticated ruminants. To date, there has been no study using high-throughput/next generation sequencing to determine the complexity of the bacterial community present in the rumen of the moose, nor has there been a comparison of moose from different geographic locations to determine the core phylotypes which exist in moose. It was hypothesized that bacterial populations would be distinct based on geographic location, and that previous studies using classical approaches [5] had underestimated diversity in the rumen of moose. In previous studies, age [9, 10] and geographic location [11] play a role in differing core microbiomes.
3.3 Methods

3.3.1 Sample Collection

A total of 17 samples (Table 1) were collected from three different geographic locations over two years. In October 2010, eight whole rumen samples were collected from moose in Vermont (VT) shot during the 2010 hunting season (October 16-21, 2010). Samples were collected by licensed hunters, during field dressing, within 2h of death and fixed frozen. Within 24 h, samples were transported to the laboratory and stored at -20°C until DNA extraction at the University of Vermont (UVM), Burlington, VT. Hunters received written and verbal instructions on sample collection to sample from well inside the rumen and to fill the container to minimize oxygen exposure. Institutional Animal Care and Use Committee (IACUC) approval was not required to collect rumen samples from licensed hunters.

Likewise, in the fall of 2011, six samples were collected from licensed hunters in Norway (NO). The Norwegian moose hunting season is much longer (September 1-October 31) than that in the United States, thus hunters may be in the field for weeks at a time. To accommodate this, whole rumen samples were collected during field dressing and immediately fixed in 70% ethanol until they were brought to the Department of Arctic Biology, University of Tromsø, Tromsø, Norway. Samples were stored at 4°C until DNA extraction at the University of Tromsø by the corresponding author. The DNA extract, containing 0.1 volume of 2M sodium acetate, then two volumes of 100% ethanol, was shipped to UVM. Once there, samples were centrifuged at 14,000g for 30 min, supernatant was then poured off, and the pellet washed with 100μl of 100% ethanol. The pellet was air dried and then suspended in 50μl of TE (Tris-EDTA, pH 8.0) buffer.
In August, 2012, three whole rumen samples were collected via esophageal tubing of sedated captive, free-range wild moose at the Moose Research Center, Soldotna, Alaska (AK) (IACUC protocol #11-021, UVM; ACUC protocol #2011-026, Department Fish and Game, Alaska). Wild moose there live in an approximately 2 square mile enclosure where they can forage naturally. Rumen samples were immediately fixed with 70% ethanol and shipped to UVM for DNA extraction. Care was taken to prevent contamination of rumen digesta sample with saliva. In all samples, approximately 50 g of whole rumen digesta sample was collected. Previously, it was shown that different methods of rumen sample collection do not affect the rumen community structure [12].

Vermont samples were collected from October 16-23, 2010 when temperatures are historically 2-12°C (low/high), however, in 2010 most areas were unusually warm with daytime highs around 21°C [13]. Norwegian samples were collected between September 26-October 6, 2011, with temperatures ranging from (low high) 6-13°C in September, (historically 3-9°C), and 0.8-5°C in October (historically 0-5°C) [13]. Alaskan samples were collected on August 31, 2012, with temperatures ranging from 9-12°C (historically 5-15°C) [13]. It is important to note that Vermont moose were sampled during summer-like temperatures, hotter than either Alaskan or Norwegian moose, despite being sampled at the latest time point in the year. In 2010, Vermont also received nearly twice the annual average rainfall [14], in 2011, Troms County, Norway received its highest annual rainfall since 1983 [13]; and in 2012, the region around Soldotna, Alaska had its highest annual rainfall in six years [13].
3.3.2 DNA extraction and quantification

Samples were fully thawed and 0.25 gram aliquots of whole contents (liquid and particle associated) were used for extraction. DNA was extracted from all 17 samples individually using the repeated bead-beating plus column (RBB+C) method [14], combined the QIAamp DNA Stool Mini Kit (QIAGEN, Maryland) and/or the Powersoil DNA Isolation Kit (MO BIO Laboratories, California). Samples were homogenized using zirconia beads for 3 min, then incubated with lysis buffer [15] at 70°C for 15 min, followed by centrifugation at 4°C for 5 min at 16,000G. This was performed twice, and the supernatant from each was combined and treated with an inhibitex tablet from the QIAGEN kit. The remainder of the DNA extraction followed the manufacturer’s instructions. Final elutions were made into 200 μl of TE (Tris-EDTA, pH 8.0) buffer, and DNA was quantified using a NanoDrop 2000C Spectrophotometer (ThermoScientific, California). PCR was performed on a C1000 ThermoCycler (Bio-Rad, California) using the iTaq kit (Bio-Rad, California) to measure the quality of DNA from the mixed sample of bulk nucleic acids. All PCR results were run on a 1% agarose gel at 100 volts for 60 min, and imaged on a ChemiDoc XRS+ gel imager (Bio-Rad, California).

3.3.3 Amplicon library preparation

The variable regions V1-V3 of the bacterial 16S rRNA were recently determined to be the overall best region to estimate species-level richness using genetic distances of 0.03 and 0.04 for cultured and uncultured bacterial sequences, while providing an optimal amplicon size of approximately 500b [16]. DNA was PCR amplified with universal
bacterial primers (IDT, California): 27F [7], (5’-AGAGTTTGATCCTGGCTCAG -3’)
and 519R [8], (5’-GWATTACCGCGGCKGCTG-3’). These primers were selected to
amplify the first three variable regions (V1-V3) of the 16S rRNA gene in bacteria,
creating an amplicon of ~500bp. The iProof High Fidelity DNA polymerase kit (Bio-Rad,
California) was used: 10µl of 5X High Fidelity buffer which includes MgCl2, 1.0µl of
10mM dNTP mix, 2.5µl each of forward and reverse primer 0.5µl of iProof DNA
polymerase and 31.5µl of ddH2O. For each sample, 2µl of DNA template were added
once the master mix had been aliquoted, to a total reaction volume of 50µl. A PCR
procedure was developed to optimize amplification, as follows: initial denaturing at
98°C for 4 min, then 34 cycles of 98°C for 10 s, 50°C for 30 s, 72° for 2 min, followed
by a final extension step of 72°C for 10 min. All PCR results were run on a 1% agarose
gel, and each sample was run in duplicate or triplicate.

The bands from each moose sample were excised from the agarose gel, combined
per sample, and purified using the QIAGEN QIAQuick Gel Extraction Kit (QIAGEN,
Maryland) according to manufacturer’s instructions. The gel-extracted DNA was re-
eluted into 30µl Buffer EB, and was quantified using a NanoDrop 2000C
Spectrophotometer (ThermoScientific, California) to a minimum required final
concentration of 20ng/µl per 20µl sample. The DNA amplicons were frozen and shipped
overnight to Molecular Research, LP (MR DNA) for Roche 454 pyrosequencing with
Titanium chemistry.
3.3.4 Sequencing analysis

Sequences were deposited online in the Sequence Read Archive (SRA) though NCBI, under study accession number SRP022590. To analyze the DNA sequencing data, the open-source computer software program MOTHUR ver.1.29 [17, 18] was used. Sequences from all three locations were processed together using the original standard flowgram format (sff) output file from the sequencer. Flow grams were denoised using “shhh.flows”, (i.e. the MOTHUR-integrated version of the PyroNoise algorithm [19]). The barcode and forward primer were removed, and sequences which contained any of the following conditions were discarded: < 300 bases, >550 bases, contained homopolymer runs >8 bases, or contained any mismatches in the barcode. After trimming, identical sequences were grouped into “unique” sequences using MOTHUR, which allowed for a comparison of total sequences, as well as those sequences which were differentiated by at least one base.

Sequences were aligned using the Needleman-Wunsch global alignment algorithm [20], 8b kmer searching, match reward +1, mismatch penalty -1, and gap open/extend penalty -2. A reference alignment for bacteria, which had been created and optimized for this type of data set in the laboratory, was used to align the candidate sequences. The alignment was then filtered to remove gap-only columns, and sequences were trimmed to a common length of 392 characters. The alignment was checked for chimeras using the MOTHUR-integrated version of the program UCHIME [21]; using abundance as a reference. These putative-chimeric sequences were classified against a Silva bacterial taxonomy [22, 23], and only those putative-chimeric sequences which had less than 80% similarity to known sequences were removed from the alignment.
The remaining aligned sequences were all classified using the Silva reference taxonomy and an 80% cutoff. Genetic distance was calculated, sequences clustered into operational taxonomic units (OTUs) at 0.03% and 0.05% genetic distance using the nearest neighbor method, and a representative sequence chosen for each cluster/OTU at each distance. Sequences were subsampled from each moose rumen sample, and diversity was compared using CHAO [24], ACE [25], Good’s Coverage [26] and the Shannon-Weiner Index [27]. Good’s Coverage, \( C = 1 - \frac{N_1}{n} \), where \( C \) is the coverage of a random sample of size \( n \), and \( N_1 \) is the numbers of “classes” observed once. A large number of “classes” observed only once will create a value of \( C \) approaching 0. Shared OTUs were generated using the make.shared command in MOthur and manual interpretation of the table. Shared OTUs were also compared using the get.coremicrobiome command.

A relaxed neighbor-joining tree was calculating using the Clearcut tree making program within MOthur, and then used to run a weighted and unweighted Unifrac within MOthur was run using random sampling. The distance file created from the weighted Unifrac was used to create a Principal Component Analysis (PCoA) and then analyzed for population differentiation using an analysis of molecular variance (AMOVA) test. Unifrac tests were also run online using FastUnifrac [28] to verify sample structure via clustering, as well as using PCoA plots (viewed in 3D with Kineimage online), P test significance, and Unifrac significance of all samples.
3.4 Results

3.4.1 Classification of taxa by sample location

The breakdown of phyla per sample is presented in Figure 1, with the statistical analysis of the sample populations as follows. The phyla Bacteroidetes, Firmicutes, and Proteobacteria, and the group “Unclassified” were found in each sample location. Overall, the phylum Bacteroidetes was dominant in the AK (69.1% of total sequences) and NO samples (40.4% of total sequences), but not from VT samples (25.7% of total sequences) (Figure 1). Within the phylum Bacteroidetes, Prevotellaceae was the dominant family in all three sample locations, with 10,522 unique sequences (33,099 total sequences) across all 17 samples. Within the Prevotellaceae family, 8,526 sequences were identified as belonging to the genus Prevotella: 5,511 sequences from AK moose (50.4% of unique sequences), 2,625 sequences from NO moose (17.3% of unique sequences), and 390 sequences (5.2% of unique) from VT. The second largest group in the phylum Bacteroidetes was the group “RC9” (family Rikenellaceae), with 495 unique sequences (1,827 total sequences) across all samples.

Bacteria belonging to the phylum Firmicutes were dominant in the VT samples (56.4% of total sequences), were the second most prevalent in NO samples (32.8% of total sequences), and third most prevalent in AK samples (9.3% of total sequences) (Figure 1). Lachnospiraceae was the most prevalent family within the phylum Firmicutes with 6,677 unique sequences (21,252 total) across all samples, followed by Ruminococcaceae with 1,303 unique sequences (3,018 total). The largest groups at the genus-level were as follows: Butyrivibrio: 457 unique sequences (1,706 total), Syntrophococcus- 186 unique sequences (514 total), Butyrivibrio-Pseudobutyribrio-
110 unique sequences (402 total), *Ruminococcus* - 60 unique sequences (155 total), *Mitsuokella* - 50 unique sequences (84 total), *Moryella* - 37 unique sequences (123 total), *Mogibacterium* - 30 unique sequences (47 total), and *Lachnospira* - 24 unique sequences (64 total).

Uncharacterized and unclassified bacteria at the phylum level represented a large proportion of sequences: 12.8% unique sequences (18.7% of total) in AK, making it the second largest group in those samples, 19.5% unique sequences (19.1% of total) in NO and 14.8% unique sequences (14.7% of total) in VT (Figure 1). Representatives of the phylum Proteobacteria were the fourth most prevalent bacteria across all 17 samples in each of the three sample locations. Within this phylum, the most prevalent genera were *Acinetobacter*: 173 unique sequences (542 total), and *Pseudomonas*: 137 unique sequences (1,247 total). The phyla Cyanobacteria and Actinobacteria were also found in each sample location, but they were much more prevalent in NO moose (2% and 1% of sequences, respectively) than in AK or VT moose (<1% and <0.4% of sequences each). The phyla Lentisphaerae, Spirochaetes, and Synergistetes, as well as the candidate division “TM7” were found in all three sample locations, with each representing <0.4% of sequences in any sample location. Fusobacteria was only found in NO samples, Acidobacteria and Chloroflexi were only found in VT samples, and candidate division “SR1” was found in NO and VT samples, with <0.4% of sequences in any sample location.

The following 68 genera were also identified, but with low frequency (i.e. <100 unique sequences): *Acetitomaculum, Acetobacter, Acidovorax, Adlercreutzia, Afipia, Anaerobiospirillum, Anaerococcus, Anaerostipes, Ancalomicrobium, Aquabacterium,

3.4.2 Statistical analysis of OTUs and clustering

Overall, a total of 109,643 sequences were generated from the 17 samples. Of these, a total of 37,831 sequences from the three Alaskan samples (AK1R-AK3R) revealed 10,936 unique sequences, 51,459 total sequences from the six Norwegian samples (NO1R-NO6R) revealed 15,211 unique sequences, and, 20,353 total sequences from the eight Vermont samples (VT1R-VT8R) revealed 7,475 unique sequences. The number of total sequences per sample (Table 1), which passed the quality control and processing steps, and which were subsequently used in the analysis, ranged from 483 to 19,583. Of the 33,622 unique sequences, 28,111 were classified to phylum (not including “Unclassified”), 21,122 down to family, and 10,642 down to genus. The unique and total
number of sequences for each sample location at the family level of classification is provided (see Supplemental Table 1).

Using a genetic distance of 3%, the 109,662 total sequences were assigned to 17,774 OTUs, of which 15,271 contained a single sequence. Owing to the wide range of sequences per sample (483 to 19,583), a random subsample of the sequences was selected, using the smallest sample size as the subsample size, and diversity indices measured. The CHAO [24] and ACE [25] richness estimators, Good’s [26] coverage, and Shannon-Weiner [27] diversity index were calculated based a 3% genetic distances for each sample (Table 2). For the entire dataset, at a 3% genetic distance, and using a subsample consisting of 483 sequences (i.e. 391 OTUs), the estimated number of OTUs which should have been observed were CHAO=3,279 and ACE=8,420, Good’s [26] coverage was 0.261, likely low due to the high proportion of OTU singletons, and Shannon-Weiner [27] index was 5.726.

No OTUs were shared across all 17 samples, using either get.coremicrobiome or a shared OTU table (Table 3) at a 3% genetic distance. Overall, only two OTUs were found at a relative abundance of greater than 1% and were found in at least 10 samples. No OTUs were present, in any abundance, in 11 or more samples. The Alaskan samples shared the greatest number of OTUs (n=92), most of which belonged to the genus *Prevotella* (Table 3). Among the NO samples (n=6), there were 8 shared OTUs, and among the VT samples (n=8) there was only 1 shared OTU (Table 3).

When comparing gender, all female samples (n=11) shared 1 OTU, and all male samples (n=6) shared 6 OTUs (Table 3). When compared by weight, only 12 samples were taken into account. Five samples were discounted because the three AK samples
only had estimated live weights, whereas the other 12 samples reported dressed carcass weight, and samples NO4R and VT4R did not have any recorded weights. The largest grouping of shared OTUs was seen in the smallest weight class, 0-100kg (NO1R, NO6R), which shared 67 OTUs, followed by the largest weight class, 301-400kg (VT5R, VT7R, VT8R) (Table 3). Weight class 101-200kg (NO2R, VT1R, VT3R) shared 26 OTUs containing 740 unique sequences, and weight class 201-300kg (NO3R, NO5R, VT2R, VT6R) shared 5 OTUs (Table 3).

Sample clustering was evaluated using weighted and unweighted FastUnifrac (Figure 2A and B, respectively). For the weighted Unifrac, there was a Unifrac significance of p=1.0e-03 (corrected), meaning that there is no probability that the branch lengths would be seen by chance, and a P-Test Significance of p=0.0 (corrected), meaning that samples were significantly clustered. In both the weighted and unweighted Unifrac, the three AK samples (AK1R-AK3R) clustered together, and six VT samples (VT3R-VT8R) clustered together (Figure 2). The remaining two VT samples and the six NO samples clustered differently between the weighted and unweighted Unifrac. In the weighted Unifrac, NO1R and NO4R clustered with the large VT clade, while the other four NO samples clustered separately with VT1R (Figure 2A). Samples NO2R, NO3R, NO5R, NO6R, and VT1R were all between calf age and approximately 3 years of age. In the unweighted Unifrac, all six NO samples clustered together, along with VT1R and VT2R (Figure 2B). The NO moose were all <4 years of age, and VT1R and 2R were 1 year and 3 years old, respectively. However, this is not exclusive, as VT3R and VT6R were aged 2 years and 3 years, respectively, yet clustered with older and heavier moose.
In both weighted and unweighted Unifrac, five out of six males clustered together excluding NO5R, the only male NO moose sample. Additionally, VT6R (female) clustered with the males both times, and VT2R (female) clustered with males in the weighted analysis. The five males and two females mentioned above were also the seven heaviest animals, with the exception of the one male from Norway which clustered with the five female NO moose samples.

According to PCoA, samples clustered most significantly by geographic location (Figure 3A), and males clustered significantly, though females did not (Figure 3B). There was not a strong clustering based on weight class (Figure 3C) or age (data not shown). According to AMOVA, when comparing diversity across different factors, it was shown that groups were statistically heterogeneous based on gender (F-statistic=2.10077, P=0.042), geographic location (Fs=4.63938, P=<0.001), and weight class (Fs=2.01592, P=0.003).

3.5 Discussion

3.5.1 Bacteroidetes:Firmicutes

Bacteroidetes was the dominant phylum in rumen samples from AK and NO, and Firmicutes was the dominant phylum in samples from VT, the latter has been previously demonstrated [6]. Firmicutes was the dominant phylum in Thompson’s gazelle, Grant’s gazelle, and eland [29]; in Norwegian reindeer [11], and in dromedary camels in Australia [30]. There are two studies on rumen bacteria from Svalbard reindeer; one study reporting Bacteroidetes to be dominant in reindeer on a winter diet [31] and the other reporting Firmicutes to be dominant in reindeer on a late summer diet [11]. This is
contrary to a another study using traditional culturing of bacteria from arctic Svalbard reindeer, which found that *Streptococcus bovis* (phylum Bacteroidetes) were more abundant in summer, along with increased starch utilization, proteolysis and lactate utilization by rumen microbes, whereas fiber digestion, cellulolysis and xylanolysis were increased during the winter, along with the cellulolytic bacteria *Butyrivibrio fibrisolvens* (phylum Firmicutes) [32]. This shift in bacterial phyla dominance reflects the change in quality and consistency of seasonal diets, as high starch diets and high fiber diets favor different rumen bacteria.

In the present study, the Lachnospiraceae represented the largest family in the phylum Firmicutes for all three sample locations, and it was the most abundant family in the rumen of VT moose. Rumen Lachnospiraceae are a family of bacteria which produce butyrate, a short-chain fatty acid, which is readily absorbed and metabolized by rumen epithelial cells [33]. Butyrate reduces the side effects of gastrointestinal inflammation, and stimulates rumen development by increasing rumen epithelial papillae growth [34]. The prevalence of Lachnospiraceae has been observed previously in studies of the Tammar wallaby, an Australian herbivorous marsupial [35], dairy cows [36], North American moose [6], and various arctic ruminants [27]. Furthermore, Lachnospiraceae was the second-most abundant bacterial family in dromedary camels in Australia [30]. In the folivorous bird, the Hoatzin, chicks had the highest proportion of Lachnospiraceae, which decreased as age increased [9]. In the present study, the observed number of Lachnospiraceae per sample were plotted against age using the statistical package JMP ver.9.0, and while there appeared to be an inverse relationship between increase in age
and decrease in the family Lachnospiraceae, the correlation was not significant ($R^2=0.20$) (data not shown).

Bacteroidetes has previously been shown to be dominant in growing ruminants [10], or ruminants transitioning from a high-fiber diet to a high-starch/high-digestibility diet [37], hibernating ground squirrels [38], in the oral cavity of pregnant women [39], and has been associated with increased fat deposition [40, 41]. Studies involving fecal transfers from pregnant women to gnotobiotic mice increased the proportion of the phyla Proteobacteria and Actinobacteria and decreased the proportion of the phyla Firmicutes and Bacteroidetes, which mirrored the change in women over the course of gestation [41]. In the present study, only one female moose, NO4R (age and weight unknown), was confirmed to have a calf, and all other females had an unknown reproductive status. Proteobacteria was not elevated in this particular sample, but it had the second highest prevalence of Actinobacteria, after sample NO2R (female, 1.5 year, 120kg dressed carcass weight).

Previously, it was shown using the G2 PhyloChip microarray that the rumen of VT moose was largely made up of bacteria in the phylum Firmicutes, followed closely by bacteria in the phylum Proteobacteria [6]. In the present study, Proteobacteria represented 4% of total sequences for NO moose, 2.5% of total sequences for AK moose, and only 1.2% of total sequences for VT moose. This discrepancy is most likely due to the bias inherent in the G2 PhyloChip, which was created using 16S rRNA oligonucleotide probes, most of which are designed to target multiple taxa [40], and which likely produced a large number of false positives [6].
3.5.2 Temperature, region, or life stage?

The samples in the present study were collected during unusually rainy seasons in all three sample locations, and unusually warm seasons in Vermont and Norway [13, 14]. Previously, it has been shown that hot and rainy summers, and not cold summers or winters, decreased average moose weight in Norway [43]. We speculate that this would cause plants to mature faster and transition more quickly during the growing season from a high proportion of starch to a high proportion of structural carbohydrates. Plant starch is more nutritious, and is likely to be associated with a higher proportion of ruminal Bacteroidetes, while structural carbohydrates (i.e. cellulose, hemicellulose, lignin), are very fibrous and are likely be associated with a higher proportion of ruminal Firmicutes [37, 44]. The unusually mild fall temperatures and heavy rainfall in Vermont may have increased the cellulose content of forage and, therefore, resulted in a high proportion of the primarily cellulolytic phylum Firmicutes in the VT moose rumen. This may explain why the VT moose had a different microbial population than the AK or NO moose. Moreover, the cooler temperatures in Alaska may have led to a higher starch content in forage, resulting in moose which had a high proportion of the phylum Bacteroidetes.

AK samples consistently clustered together and shared the greatest number of OTUs and unique sequences. This is unsurprising, as these moose cohabited the same 2 square mile enclosure for years, were likely consuming a similar diet, and were roughly the same age. However, the VT and NO samples each clustered separately with a large amount of consistency, which was not observed among age or weight groups across all sample locations. This suggests that geographic location and, thus, available forage, plays a large role in defining the core microbiome in the three isolated populations in the
present study [11]. Age/life stage may play a larger role in defining the core microbiome within a sample location and account for individual variation [6, 9, 41]. While males did cluster on PCoA graphs and Unifrac, and shared six OTUs, this was not seen in females, potentially due to their different reproductive stages. Although there was significant clustering of females based on weight/age using Unifrac, it was not corroborated using PCoA, and they shared only one OTU. This discrepancy could be due to the low number of samples per age/weight group.

3.6 Conclusion

The present study found that samples taken from Vermont (USA), Alaska (USA), and Norway differ from each other in terms of phylogenetic diversity with respect to geographical location, gender, and weight of the host animal. The observed clustering trends, between moose populations, are likely due to different quality of diet between populations, which may be due to differing location, climate, and sampling season between the three groups.

3.7 Availability of Supporting Data

The data set supporting the results of this article is available in the Sequence Read Archive (SRA), available through NCBI [study accession number SRP022590].

3.8 Competing Interests

The authors declare no competing interests.
3.9 Authors’ Contributions

SLI performed all laboratory procedures, excluding Roche 454 pyrosequencing, analyzed the dataset, and wrote the manuscript. ADW conceived of the project, facilitated Norwegian sample collection, funded all work and edited the manuscript.

3.10 Acknowledgments

The authors would like to acknowledge: the Vermont Fish and Wildlife Dept. for sample collection logistics; Terry Clifford, Archie Foster, Lenny Gerardi, Ralph Loomis, Beth and John Mayer, and Rob Whitcomb for collection of Vermont moose samples; Dr. Even Jørgensen, University of Tromsø, and Dr. Helge K. Johnsen, University of Tromsø, for collection of Norwegian moose samples; Dr. Monica A. Sunset, University of Tromsø, for facilitating sample collection and storage, as well as for providing DNA extraction materials for the Norwegian samples; Dr. John Crouse and Dr. Kimberlee Beckmen, both of the Alaska Dept. of Fish & Game for collection of Alaskan moose samples; and Dr. Benoit St-Pierre, University of Vermont for assistance with MOTHUR and Perl programming.

3.11 References


18. MOTHUR [www.mothur.org].


3.12 Figures

Figure 3-1 A comparison of sequences per sample, using the Silva bacterial reference taxonomy for classification.
Alaska (AK): 37,831 total; Norway (NO): 51,459 total; Vermont (VT): 20,353 total; overall: 109,643 total.
Figure 3-2 FastUnifrac clustering of all 17 samples. A) weighted, non-normalized, and B) unweighted.
Figure 3-3 PCoA graphs colored by variable.
A) geographic location: red=AK, blue=NO, yellow=VT, B) gender: red=female, blue=male, C) weight class: pink=not available, red=0-100 kg, light blue=101-200 kg, yellow=201-300 kg, green=301-400 kg, and purple=400+ kg for live weights.
### 3.13 Tables

**Table 3-1 Metadata of the 17 samples from Alaska (AK), Norway (NO), and Vermont (VT).**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sample Site</th>
<th>Collection m-d-y</th>
<th>Sex</th>
<th>Weight (kg)(^1)</th>
<th>Approx. age(^2)</th>
<th># Seqs used for analysis(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK-1R</td>
<td>Soldotna, AK 60°29'N 151°4'W</td>
<td>08-31-12</td>
<td>F</td>
<td>485 (live wt.)</td>
<td>10y 3mo</td>
<td>5,695</td>
</tr>
<tr>
<td>AK-2R</td>
<td>“ ”</td>
<td>08-31-12</td>
<td>F</td>
<td>487 (live wt.)</td>
<td>10y 3mo</td>
<td>12,550</td>
</tr>
<tr>
<td>AK-3R</td>
<td>“ ”</td>
<td>08-31-12</td>
<td>F</td>
<td>506 (live wt.)</td>
<td>11y 3mo</td>
<td>19,586</td>
</tr>
<tr>
<td>NO-1R</td>
<td>Troms County, NO 69°N 20°E</td>
<td>09-26-11</td>
<td>F</td>
<td>61 (dressed)</td>
<td>Calf</td>
<td>6,028</td>
</tr>
<tr>
<td>NO-2R</td>
<td>“ ”</td>
<td>09-26-11</td>
<td>F</td>
<td>120 (dressed)</td>
<td>1.5y</td>
<td>18,543</td>
</tr>
<tr>
<td>NO-3R</td>
<td>“ ”</td>
<td>09-26-11</td>
<td>F</td>
<td>211 (dressed)</td>
<td>&gt;4.5y</td>
<td>3,122</td>
</tr>
<tr>
<td>NO-4R</td>
<td>“ ”</td>
<td>09-30-11</td>
<td>F</td>
<td>N/A</td>
<td>N/A(^4)</td>
<td>3,037</td>
</tr>
<tr>
<td>NO-5R</td>
<td>“ ”</td>
<td>09-27-11</td>
<td>M</td>
<td>290 (dressed)</td>
<td>2-3y</td>
<td>17,624</td>
</tr>
<tr>
<td>NO-6R</td>
<td>Andørja, Ibestad, NO 68°N 17°E</td>
<td>10-08-11</td>
<td>F</td>
<td>77 (dressed)</td>
<td>Calf, 6-8 mo.</td>
<td>3,105</td>
</tr>
<tr>
<td>VT-1R</td>
<td>Averill, VT 44°59'N, 71°42'W</td>
<td>10-16-10</td>
<td>F</td>
<td>185 (dressed)</td>
<td>1y</td>
<td>3,484</td>
</tr>
<tr>
<td>VT-2R</td>
<td>East Haven, VT 44°39'N, 71°53'W</td>
<td>10-16-10</td>
<td>F</td>
<td>244.6 (dressed)</td>
<td>3y</td>
<td>1,744</td>
</tr>
<tr>
<td>VT-3R</td>
<td>North Danville, VT 44°27'N, 72°5'W</td>
<td>10-17-10</td>
<td>M</td>
<td>186.4 (dressed)</td>
<td>2y</td>
<td>3,405</td>
</tr>
<tr>
<td>VT-4R</td>
<td>Canaan, VT 44°59'N, 71°32'W</td>
<td>10-16-10</td>
<td>M</td>
<td>N/A</td>
<td>N/A</td>
<td>3,679</td>
</tr>
<tr>
<td>VT-5R</td>
<td>Ferdinand, VT 44°42.7'N, 71°45.19'W</td>
<td>10-20-10</td>
<td>M</td>
<td>319.1 (dressed)</td>
<td>4y</td>
<td>3,261</td>
</tr>
<tr>
<td>VT-6R</td>
<td>Brighton, VT 44°48.3'N, 71°51.3'W</td>
<td>10-23-10</td>
<td>F</td>
<td>259.6 (dressed)</td>
<td>3y</td>
<td>483</td>
</tr>
<tr>
<td>VT-7R</td>
<td>Walden, VT 44°26.3'N, 72°13.4'W</td>
<td>10-23-10</td>
<td>M</td>
<td>301.3 (dressed)</td>
<td>4y</td>
<td>2,305</td>
</tr>
<tr>
<td>VT-8R</td>
<td>Wheelock, VT 44°35.28'N, 72°5.33'W</td>
<td>10-23-10</td>
<td>M</td>
<td>405.5 (dressed)</td>
<td>8y</td>
<td>1,992</td>
</tr>
</tbody>
</table>

\(^1\)AK moose were weighed live, VT and NO moose carcasses were field dressed and weighed at reporting stations. \(^2\)AK moose had birthdates, VT and NO moose age was estimated using wear on teeth. \(^3\)Unique sequences = 33,622 \(^4\)Moose had a bull calf.
Table 3-2 Statistical measures using a subset of sequences per sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th># OTUs in sample</th>
<th>CHAO</th>
<th>ACE</th>
<th>Good’s Coverage</th>
<th>Shannon-Weiner Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK_1R</td>
<td>962</td>
<td>6,012</td>
<td>11,389</td>
<td>0.576</td>
<td>5.700</td>
</tr>
<tr>
<td>AK_2R</td>
<td>1,707</td>
<td>9,846</td>
<td>22,158</td>
<td>0.554</td>
<td>6.523</td>
</tr>
<tr>
<td>AK_3R</td>
<td>2,374</td>
<td>15,861</td>
<td>45,718</td>
<td>0.642</td>
<td>5.787</td>
</tr>
<tr>
<td>NO_1R</td>
<td>1,329</td>
<td>7,737</td>
<td>20,686</td>
<td>0.381</td>
<td>6.813</td>
</tr>
<tr>
<td>NO_2R</td>
<td>2,926</td>
<td>18,244</td>
<td>42,501</td>
<td>0.538</td>
<td>7.047</td>
</tr>
<tr>
<td>NO_3R</td>
<td>563</td>
<td>3,253</td>
<td>3,792</td>
<td>0.380</td>
<td>5.779</td>
</tr>
<tr>
<td>NO_4R</td>
<td>730</td>
<td>4,459</td>
<td>10,654</td>
<td>0.327</td>
<td>6.364</td>
</tr>
<tr>
<td>NO_5R</td>
<td>3,554</td>
<td>25,415</td>
<td>62,870</td>
<td>0.409</td>
<td>7.654</td>
</tr>
<tr>
<td>NO_6R</td>
<td>774</td>
<td>6,318</td>
<td>18,613</td>
<td>0.273</td>
<td>6.433</td>
</tr>
<tr>
<td>VT_1R</td>
<td>987</td>
<td>6,630</td>
<td>15,544</td>
<td>0.279</td>
<td>6.690</td>
</tr>
<tr>
<td>VT_2R</td>
<td>539</td>
<td>3,179</td>
<td>3,393</td>
<td>0.268</td>
<td>6.149</td>
</tr>
<tr>
<td>VT_3R</td>
<td>992</td>
<td>6,230</td>
<td>15,364</td>
<td>0.282</td>
<td>6.765</td>
</tr>
<tr>
<td>VT_4R</td>
<td>1,106</td>
<td>7,430</td>
<td>12,943</td>
<td>0.294</td>
<td>6.77</td>
</tr>
<tr>
<td>VT_5R</td>
<td>958</td>
<td>5,776</td>
<td>14,374</td>
<td>0.307</td>
<td>6.649</td>
</tr>
<tr>
<td>VT_6R</td>
<td>148</td>
<td>2,122</td>
<td>3,450</td>
<td>0.113</td>
<td>4.957</td>
</tr>
<tr>
<td>VT_7R</td>
<td>669</td>
<td>4,407</td>
<td>8,443</td>
<td>0.303</td>
<td>6.335</td>
</tr>
<tr>
<td>VT_8R</td>
<td>686</td>
<td>5,607</td>
<td>5,544</td>
<td>0.185</td>
<td>6.471</td>
</tr>
</tbody>
</table>

Table 3-3 The number of shared OTUs across different samples at a 3% genetic distance generated using a shared OTU table in MOTHUR.

<table>
<thead>
<tr>
<th>Samples Compared</th>
<th>OTUs Shared, 3% distance</th>
<th>Shared Unique Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK samples (n=3)</td>
<td>92</td>
<td>4,510</td>
</tr>
<tr>
<td>NO samples (n=6)</td>
<td>8</td>
<td>380</td>
</tr>
<tr>
<td>VT samples (n=8)</td>
<td>1</td>
<td>136</td>
</tr>
<tr>
<td>All 17 samples</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>All Females (n=11)</td>
<td>1</td>
<td>114</td>
</tr>
<tr>
<td>All Males (n=6)</td>
<td>6</td>
<td>314</td>
</tr>
<tr>
<td>Weight: 0-100 kg (NO1R, NO6R)</td>
<td>67</td>
<td>433</td>
</tr>
<tr>
<td>Weight: 101-200 kg (NO2R, VT1R, VT3R)</td>
<td>26</td>
<td>740</td>
</tr>
<tr>
<td>Weight: 201-300 kg (NO3R, NO5R, VT2R, VT6R)</td>
<td>5</td>
<td>417</td>
</tr>
<tr>
<td>Weight: 301-400 kg (VT5R, VT7R, VT8R)</td>
<td>28</td>
<td>276</td>
</tr>
</tbody>
</table>
CHAPTER 4  DESIGN AND VALIDATION OF FOUR NEW PRIMERS FOR NEXT-GENERATION SEQUENCING TO TARGET THE 18S rRNA GENE OF GASTROINTESTINAL CILIATE PROTOZOA.

Suzanne L. Ishaq and André-Denis G. Wright

Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, 570 Main St., Burlington, Vermont 05401

Contact: Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax) 1-802-656-8196.

Keywords: Alaska/Entodiniomorphida/ciliophora/moose/protist/rumen/Vestibuliferida

4.1 Abstract

Four new primers and one published primer were used to PCR amplify hyper-variable regions within the protozoal 18S rRNA gene to determine which primer pair provided the best identification and statistical analysis. PCR amplicons of 394 to 498 bases were generated from three primer sets, sequenced using Roche 454 pyrosequencing with Titanium, and analyzed using the BLAST (NCBI) database and MOTHUR ver. 1.29. The protozoal diversity of rumen contents from moose in Alaska was assessed. In the present study, primer set 1, P-SSU-316F + GIC758R (amplicon = 482 bases) gave the best representation of diversity using BLAST classification, and amplified *Entodinium simplex* and *Ostracodinium* spp., which were not amplified by the other two primer sets. Primer set 2, GIC1080F + GIC1578R (amplicon = 498 bases), had similar BLAST results and a slightly higher percentage of sequences that identified with a higher sequence identity. Primer sets 1 and 2 are recommended for use in ruminants. However, primer set 1 may be inadequate to determine protozoal diversity in non-ruminants. Amplicons created by primer set 1 were indistinguishable for certain species within the genera *Bandia*, *Blepharocorys*, *Polycosta*, *Tetratoxum*, or between *Hemiprorodon gymnoprosthium* and *Prorodonopsis coli*, none of which are normally found in the rumen.
4.2 Introduction

Rumen ciliate protozoa represent important functional members of the rumen environment, as most have some cellulolytic or amylolytic abilities (1–3). Most studies of rumen ciliate protozoa are performed using microscopy and traditional culturing techniques (2, 4–10), quantitative PCR (11, 12), denaturing gradient gel electrophoresis (13), and full-length 18S rRNA clone libraries (13, 14). A few studies of rumen ciliate protozoa use high-throughput sequencing, although primer selection remains a problem, as some studies use universal eukaryotic primers, primers which target only one ciliate protozoa signature region, or primers which produce unsuitable long amplicons for current high-throughput technology (15–20).

The 18S rRNA gene ranges from 1.5 kb to over 4.5 kb (21), and in rumen ciliate protozoa it is generally 1.5 kb to 1.8 kb in length. Like the 16S rRNA gene of prokaryotes, the 18S rRNA gene of eukaryotes has nine hyper-variable regions (V1–V9) which can be used for genus/species identification. Four gut ciliate signature regions exist within the rumen protozoal 18S rRNA gene, which represent areas of high variability that can improve identification down to species level (22,23). Signature region 1 occurs between 440–460bp (within V3), signature region 2 occurs between 590-620bp (between V3 and V4), signature region 3 occurs between 1220–1260bp (within V6), and signature region 4 occurs between 1560–1580bp (after V8) (Figure 1). Additionally, rumen ciliate protozoa have a slightly different 18S rRNA secondary structure from non-rumen ciliates, in that rumen protozoa are missing helix E23–5 from the V4 region and other helices in the region are shorter (21-23). Previously, the V9 region (15) or the V5–V7 regions (13) have been amplified for phylogenetic analysis.
using high-throughput techniques. Ciliated protozoa found in the gastrointestinal tract of animals belong to the phylum Ciliophora, class Litostomatea, subclass Trichostomatia, and the orders Entodiniomorphia and Vestibuliferida. The vast majority of rumen ciliated protozoa belong to the Ophryoscolecidae, the largest family (both in numbers of species and genera) within the Entodiniomorphia.

In the present study, four new primers were designed to specifically target conserved 18S rRNA gene regions for ciliate protozoa, which are normally found in the gastrointestinal tract of herbivores. Using our protocol, these primers did not amplify other eukaryotic, bacterial, or archaeal species, or non-ciliated protozoa which are not normally found in a healthy gastrointestinal tract environment. The strategy for the pairing of the forward and reverse primers was to create amplicons which included at least one of the four signature regions of the rumen ciliate protozoal 18S rRNA gene. Current limitations of the Roche 454 and MiSeq ver. 3.0 (with 2 x 300 base paired ends) platforms, along with few reliable conserved regions, exclusive to ciliate protozoa, prevent the inclusion of all four signature regions, which was possible when the full 18S rRNA gene was sequenced using Sanger sequencing technology.

The first objective of the present study were to test these primer pairs on rumen samples from the North American moose (*Alces alces*) to determine their suitability and validity for rumen ciliate identification. The second objective was to compare the primer sets using CHAO, ACE, Shannon-Weaver, Inverse Simpson, Good’s Coverage, and Unifrac in order to make a recommendation of the most suitable primer set for rumen protozoal 18S rRNA gene amplification.
4.3 Materials and Methods

4.3.1 Sample collection and DNA extraction

On August 31, 2012, whole rumen samples were collected via esophageal tubing of three captive, free-range wild moose at the Moose Research Center, Soldotna, Alaska (AK) (IACUC protocol #11-021, University of Vermont; ACUC protocol #2011-026, Department Fish and Game, Alaska). All three moose were females between 10-11 years of age. Rumen samples were mixed with 70% ethanol and shipped to the University of Vermont (Burlington, Vermont, USA) where they were stored at 4°C. To confirm the sequencing results, all three moose samples were visual inspected using light microscopy to identify genera of rumen ciliates.

To extract DNA, a 5 ml aliquot of whole rumen contents in ethanol was centrifuged for 5 min at 16,000 x G, and ethanol was removed by pouring off the liquid fraction. From the remaining whole contents of all the samples, 0.25 g aliquots of whole contents (liquid and particle associated) were used for extraction. DNA was extracted from the three rumen samples using the repeated bead-beating plus column (RBB+C) method (24), combined with the QIAamp DNA Stool Mini Kit (QIAGEN, Maryland). The final elutions were made using 200 μl of TE buffer (1M Tris-HCl, 0.5M EDTA, pH 8.0), and eluted DNA was quantified using a NanoDrop 2000C Spectrophotometer (ThermoScientific, California).

4.3.2 Primer design

The new forward and reverse primers were designed to target signature regions unique to gastrointestinal tract ciliate protozoa within the 18S rRNA gene. Protozoal 18S
rRNA gene reference alignments were created and used to select areas, which were highly conserved among the rumen ciliate protozoa. Four conserved regions were selected, and a potential primer sequence identified from each of those regions. The four new primers were given the prefix “GIC” for gastrointestinal ciliates, and are as shown in Table 1 with specifications. Along with a previously described rumen protozoal primer, P-SSU-316F (5’-GCTTTCGWTGGTAGTGTATT-3’) (12), primer sequences were compared to known sequences of gastrointestinal ciliates in GenBank (NCBI) (Table 2) to determine specificity to amplify only gastrointestinal tract ciliate protozoa.

Primer set 1, P-SSU-316F (12) and GIC758R, created an amplicon of 482 bases (primers not included) which encompassed variable regions V3-V4, and rumen ciliate signature regions 1-2 (Figure 1). Primer set 2, GIC1080F and GIC1578R, created an amplicon of 498b, which encompassed V6-V8 and signature regions 3-4 (Figure 1). Primer set 3, GIC1184F and GIC1578R, created an amplicon of 394b, which also encompassed V6-V8, and rumen ciliate signature regions 3-4 (Figure 1).

4.3.3 PCR amplification

The Phusion High-Fidelity DNA polymerase kit (ThermoScientific, Massachusetts) was used for PCR: 10 µl of 5X High Fidelity buffer (including MgCl2), 1.0 µl of 10 mM dNTP mix, 2.5 µl each of forward and reverse primer at 10 mM concentration, 0.5 µl of Phusion DNA polymerase (2U/µl), and 31.5 µl of ddH2O. DNA templates (2 µL of 10-50ng/µL concentration) were added once the master mix had been aliquoted, to a total reaction volume of 50 µl. The PCR protocol was adapted from previous literature (12); 94°C for 4 min hot start, followed by 30 cycles of 94°C for 30 s,
55°C for 30 s, and 72°C for 1 min, with a final extension of 72°C for 6 min on the last cycle. All PCR results were run on a 1% agarose gel at 100 volts for 60 min, and imaged on a ChemiDoc XRS+ gel imager (Bio-Rad, California).

DNA bands of the correct amplicon size were excised out of the agarose gel for DNA purification, using the QIAQuick Gel Extraction Kit (QIAGEN, Maryland) according to the manufacturer’s instructions. For each primer set, all gel bands from each of the three moose samples were filtered through the same column. The gel–extracted DNA from each primer set was quantified using a NanoDrop 2000C Spectrophotometer (ThermoScientific, California), at a minimum final concentration of 20 ng/µl, at a volume of 20 µl. The DNA amplicons from the three test primer sets were frozen and shipped overnight to Molecular Research DNA (MR DNA), Shallowater, Texas, USA, for Roche 454 pyrosequencing with Titanium chemistry.

4.3.4 Sequence analysis

Sequences were deposited online in the Sequence Read Archive (SRA) though NCBI, under study accession number SRP034591. To analyze the DNA sequencing data, the open-source computer software program MOTHUR ver.1.29 (25) was used. Sequences from all three primer sets were processed independently using the original standard flowgram format (sff) output file from the sequencer. Flow grams were denoised using the MOTHUR-integrated version of the PyroNoise algorithm (26), which also creates phylotypes, which is a unique sequence representing multiple identical sequences. Unique sequences are not equivalent to singletons, which are operational taxonomic units (OTUs) containing a single sequence. The barcode and primer
sequences were removed, and sequences which contained any of the following conditions were discarded: <400 bases for primer set 1 and 2 or <375 bases for primer set 3, >500 bases, homopolymer runs >8 bases, or any mismatches in the barcode.

To determine the genetic distance cutoffs for species-level comparisons, 51 full-length 18S rRNA valid protozoal reference sequences were obtained from NCBI (Table 2, Supplemental Figure 4). These reference sequences were then trimmed using the three test primer sets as trim points. BLAST sequences were manually aligned and pairwise distances were calculated using PHYLIP (ver. 3.69) using a Kimura 2-parameter model (Table 3). A total of 51 pairwise species (within genus) and 2,926 pairwise distances between genera were compared using validly recognized species to determine genetic distances.

Sequences were aligned using the Needleman-Wunsch global alignment algorithm (27), 8 base kmer searching, match reward +1, mismatch penalty -1, and gap open/extend penalty -2. An 18S rRNA gene reference alignment, featuring rumen and non-rumen ciliate protozoal sequences downloaded from NCBI, was created in the laboratory to provide a better alignment of candidate sequences. The reference alignment contained 219 full-length 18S rRNA sequences for all available gastrointestinal tract (rumen, forestomach, cecum, and colon) ciliates sequences, as well as non-ruminant ciliates and non-ciliate protozoal sequences from BLAST. The reference alignment contained all 51 sequences previously used to determine genetic distance cutoffs. The candidate alignment was then filtered to remove gap-only columns, and any sequences which would not align (<10 sequences per data set). The candidate alignment was checked for chimeras using the MOTHUR-integrated version of the program UCHIME
(28); using the ciliate 18S rRNA gene sequence reference alignment which was created in our laboratory. To determine the specificity of the primers, unique sequences were classified using BLAST.

Sequences from the three primer sets were trimmed to a uniform length per library (Table 3), and clustered using the Nearest Neighbor method to determine observed number of OTUs per library. Libraries were then subsampled equal to the smallest library (n=424 sequences per library), and Shannon-Weaver Diversity index (29), Good’s Coverage (30), Inverse Simpson (31), CHAO (32), and ACE (33) were calculated for each library based on the recommended genetic distance. Like Simpson’s Diversity, Inverse Simpson measures number and abundance of species. However, it weights rare species lower than Simpson to prevent a dramatic increase in diversity with the addition of rare species. Additionally, using MOTHUR, relaxed neighbor-joining trees were created from trimmed sequences (375 bases) using CLEARCUT, and trees were clustered using weighted and unweighted Unifrac as a measure of similarity of abundance and structure between libraries.

4.4 Results

4.4.1 Primer set 1: P-SSU-316F + GIC758R

A total of 12,326 sequences passed quality assurance measures and were used for sequence analysis. Of these, 769 sequences were unique (Table 3). Using aligned sequences of valid protozoa to generate pairwise genetic distances (see Supplemental Figure 1), the species and genus-level cutoffs were determined to be 0.036 and 0.087, respectively. So, a 4% species-level cutoff and a 9% genus-level cutoff were comparable
to cutoffs for near full-length gene sequences. Sequences were trimmed to 450 bases, and various diversity indices calculated for the data set. At a 4% species-level cutoff, 48 species-level OTUs were observed, and at 9% genus-level cutoff, 15 genus-level OTUs were observed (Table 3). ACE, CHAO, Shannon-Weaver and Inverse Simpson values can be found in Table 3.

Using BLAST, the most prevalent taxon was *Polyplastron multivesiculatum*, which represented nearly 60% of the unique sequences, followed by the genus *Entodinium*, which represented just over 20% of unique sequences (Figure 2). The most prevalent species of *Entodinium* were *Ent. furca dilobum* (7% of unique sequences) and *Ent. nanellum* (5% of unique sequences). *Epidinium caudatum* represented 5% of unique sequences (Figure 2). Primer set 1 amplified *Ent. simplex*, *Ostracodinium gracile*, and other *Ostracodinium* spp., which were not amplified by other primer sets. The percent identity to known sequences ranged from 95-100% for primer set 1, with 66% of sequences having a 99% identity to a known sequence in BLAST (Figure 3). However, there were six sequences that had a 99-100% identity on only 93-95% of the query sequence. The average percent identity to *Ostracodinium gracile* was 98.2% (range 97-99%). There were 21 unique (63 total) sequences, which had <96% identity to a known sequence.

During the calculation of genetic distances using pairwise comparisons, it was noted that the amplicon created by primer set 1 could not differentiate between *Blepharocorys microcorys* (AB794975) and *Blepharocorys uncinata* (AB530162), between *Hemiprorodon gymnoprosthium* (AB795028) and *Prorodonopsis coli* (AB795029), between *Tetratoxum excavatum* (AB794971) and *Tetratoxum parvum*.
(AB794969), between Bandia deveneyi (AY380823) and Bandia smalesae (AF298822), and between Polycosta roundi (AF298819) and Polycosta turniae (AF298818).

4.4.2 Primer set 2: GIC1080F + GIC1578R

A total of 6,070 sequences for this primer set passed quality assurance measures and were used for sequence analysis. Of these, 697 sequences were unique (Table 3). Using aligned sequences of valid protozoa to generate pairwise genetic distances (see Supplemental Figure 2), the species and genus-level cutoffs were determined to be 0.039 and 0.079, respectively. For statistical analysis, sequences were trimmed to a uniform length of 450 bases. At a 4% species-level genetic distance cutoff, 25 species-level OTUs were observed, and at 8% genus-level cutoff, 14 genus-level OTUs were observed (Table 3). ACE, CHAO, Shannon-Weaver and Inverse Simpson values can be found in Table 3.

Sequences from this primer set represented over 60% Polylastron multivesiculatum (Figure 2). The next predominant genus was Entodinium with 30% of unique sequences and, of that, Ent. nanellum represented approximately two-thirds of the genus (20% of unique sequences) (Figure 2). Diploplastron affine and Epidinium caudatum were the third most prevalent taxa, with 5% of unique sequences each. The percent identity to known sequences ranged from 94-100% for primer set 2, with 67% of sequences having a 99% identity to a known sequence in BLAST (Figure 3). Only 1 unique (9 total) sequence had <96% identity to known sequences.
4.4.3 Primer set 3: GIC1184F + GIC1578R

A total of 8,265 total sequences passed quality assurance measures and were used for sequence analysis. Of these, 424 sequences were unique and used for BLAST (Table 3). Using aligned sequences of valid protozoal to generate pairwise genetic distances (see Supplemental Figure 3), the species and genus-level cutoffs were determined to be 0.042 and 0.096, respectively. For statistical analysis, sequences were trimmed to a uniform length of 375 bases. At a 4% species-level genetic distance cutoff, 20 species-level OTUs were observed, and at 10% genus-level cutoff, 12 genus-level OTUs were observed (Table 3). ACE, CHAO, Shannon-Weaver and Inverse Simpson values can be found in Table 3.

Over 75% of unique sequences in the primer set 3 dataset were classified as Polyplastron multivesiculatum using BLAST (Figure 2). The next predominant genus was Entodinium, representing approximately 15% of the unique sequences (Figure 2). Epidinium was the third most prevalent genus, with <5% of the unique sequences. The percent identity to known sequences ranged from 95-99% for primer set 3, with 76% of sequences having a 98% identity to a known sequence in GenBank (Figure 3). There were 2 unique sequences, which had <96% identity to a publically available sequence.

4.4.4 Comparison of the three primer sets

Primer set 1 (P-SSU-316F + GIC758R) had the highest number of observed OTUs, as well as the highest ACE, CHAO, Shannon-Weaver and Inverse Simpson values of all three primer sets, indicating the highest amount of diversity of the three amplicon libraries (Table 3). Primer set 1 had the lowest Good’s coverage (0.95). Primer set 2
(GIC1080F + GIC1578R) had the second largest number of observed OTUs, and second largest CHAO. However, primer set 2 had the lowest ACE, Shannon-Weaver, and Inverse Simpson values, indicating a low amount of diversity (Table 3). Primer set 3 had the lowest number of observed OTUs, as well as the lowest CHAO estimate (Table 3). However, primer set 3 had the second highest ACE, Shannon-Weaver, and Inverse Simpson values (Table 3). Weighted and unweighted Unifrac analyses were also run using MOTHUR. Primer sets were not significantly different based on weighted (0.96 – 1.0), or unweighted (0.98 – 1.0, p<0.001) Unifrac. There was no correlation between sequence length and % identity to known sequences in BLAST for any of the three data sets.

Genera of rumen ciliates were confirmed using light microscopy. Various species of Entodinium, as well as Polyploastron multivesiculatum and Epidinium cattanei were found in abundance in all three samples. Isotricha were found in two samples, and Ostracodinium was found in one sample.

4.5 Discussion

This study validated three primer sets for the amplification of gastrointestinal tract ciliate protozoa for use in high-throughput sequencing, as well as determined the diversity of moose rumen protozoa from Alaska, USA. All three test primer sets were able to amplify 18S rRNA protozoal sequences. The 18S rRNA gene is highly conserved among eukaryotes, and finding potential primer sites, which are specific to certain taxa can be challenging. Using the new primers reported in the present study, under the same amplification parameters (i.e. PCR annealing temperature, removal of short amplicons),
only 18S rRNA genes from gastrointestinal tract (rumen, forestomach, cecum and colon) ciliate protozoa should be targeted for amplification.

Previously used primers for high-throughput sequencing were often universal eukaryotic primers (15-19), or primers which targeted only one signature region for the ciliate protozoa (18, 20). In several studies, this resulted in sequences which were not ciliate or protozoan in nature (17, 19), or which were too short (<200 bases), thereby increasing the risk for misidentification (18) given the overall high degree of conservation of the 18S gene across eukaryotic taxa. In the present study, no non-protozoal eukaryotic sequences (i.e. plant, fungal or host DNA) were amplified.

Previously, using classical microbiology, Dehority (6) identified just five species of protozoa from the rumen of Alaskan moose, Sládeček (4) identified four species from moose, and Westerling (34) identified just two species from moose in Finland. In the present study, 12 species representing 7 genera were found across the three Alaskan moose. Previously, between 16 to 24 rumen ciliate species, represented by 1 to 9 genera, were found in in studies of wild reindeer (9, 35–39), wild musk ox (Ovibos moschatus) (6), wildebeest (Connochaetes spp.) (40), Kafue lechwe antelope (Kobus leche kafuensis) (41), Sassaby antelope (Damaliscus lunatus) (10), and tsessebe (Damaliscus lunatus lunatus) (42).

Only one study exists, which investigated the rumen protozoa from three bull-moose in Alaska, USA, using culturing and microscopy techniques (6). Previously, Entodinium alces and Entodinium exiguum were reported to be the two dominant species in moose rumen contents, whereas Entodinium dubardi, Entodinium simplex, and Entodinium longinucleatum were present in one moose (6). Additionally, Entodinium
'dubardi' and *Epidinium caudatum* were identified in moose rumen samples from Finnish Lapland (34), and *Entodinium dubardi, Entodinium simplex, Ostracodinium obtusum,* and *Epidinium ecaudatum* were isolated from moose in Slovakia (4). *Eudiplodinium neglectum* was also first identified in a moose from Canada (5).

Based on the previous studies which identified the rumen ciliate protozoa in the moose, it was surprising that *Polyplastron* are the dominant species in the present study. *Polyplastron* produce xylanases, carboxymethylcellulases, and various other endoglucanases, which digest fiber in the rumen. While the presence of *Polyplastron* was validated using light microscopy in the present study, a large number of sequences related to *Polyplastron* may be explained by rRNA copy numbers in ciliates, which could be highly variable from one species to another. Also, there is a lack of publically available sequences for the rumen ciliates, especially closely related genera to *Polyplastron*, such as *Elytroplastron* and *Eudiplodinium*. This is also true of *Entodinium alces, Entodinium exiguum, Ostracodinium obtusum, Epidinium ecaudatum,* and *Eudiplodinium neglectum,* all of which were previously found in moose (4–6), but no representative sequences exist for these species. This makes identification at a species level very difficult until additional sequences from all described species are elucidated.

There were 74 total sequences (24 unique sequences), which had <96% identity to publically available sequences. Given that the genetic distance between *Epidinium caudatum* and *Epidinium ecaudatum* is 1.3%, some sequences in the present analysis with genetic distances between 1.0-1.5% to *Epi. caudatum* may represent other species of *Epidinium*, such as *Epidinium cattanei,* which was observed under light microscopy. Similarly, given that *Polyplastron multivesiculatum* has 98% sequence identity to
*Ostracodinium gracile* and *Ostracodinium clipoleum*, the large number of sequences which have <98% identity to *P. multivesiculatum*, may in fact represent other closely related species or genera. As we stated previously, more representative sequences from other rumen ciliates, such as *Epi. cattanei*, are needed to confirm these interpretations of the data.

Based on the analysis in the present study, the primer set which gave the best representation of diversity using BLAST was primer set 1, P-SSU-316F + GIC758R. Primer set 1 amplified *Entodinium simplex* and *Ostracodinium* spp., which were not amplified by either of the other two primer sets. However, primer set 2 had a slightly higher percentage of sequences classified at a higher % identity in BLAST than primer set 1. Both primer sets produced an amplicon of greater than 400 bases. Primer set 1 (P-SSU-316F + GIC758R) had the highest number of observed OTUs, as well as the highest ACE, CHAO, Shannon-Weaver and Inverse Simpson values of all three primer sets, indicating the highest amount of diversity of the three amplicon libraries. Primer set 3 had the second highest ACE, Shannon-Weaver, and Inverse Simpson values. While primer sets 2 and 3 target the same variable and signature regions, primer set 2 spans a larger area of conserved region, which may account for the lower estimated diversity than primer set 3.

It is important to note that primer set 1 (P-SSU-316F + GIC758R) produced amplicons which could not differentiate between the pairs *Blepharocorys microcorys* and *Blepharocorys uncinata*, *Hemiprorodon gymnoprosthium* and *Prorodonopsis coli*, *Tetratoxum excavatum* and *Tetratoxum parvum*, *Bandia deveneyi* and *Bandia smalesae*, and *Polycosta roundi* and *Polycosta turniae*. The aforementioned *Blepharocorys* spp.
Hemiproodon gymnoprosthium (44), Prorodonopsiis coli (45), and Tetratoxum spp. (46) have previously been found in the hind gut of the horse, and Bandia deveneyi, B. smalesae, Polycota roundi, and P. turniae in Australian marsupials (47). Since these ciliate species mainly occur in non-ruminants, either P-SSU-316F + GIC758R or GIC1080F + GIC1578R could be used in ruminants, but only GIC1080F + GIC1578R should be used in non-ruminants. However, in order to bring about standardization of the amplification of rumen ciliates and to better enable comparison across studies, GIC1080F + GIC1578R seems to be the better choice for a general gut ciliate primer set.

4.6 Acknowledgements

The authors would like to acknowledge Dr. John Crouse and Dr. Kimberlee Beckmen of the Alaskan Department of Fish and Game for their assistance in rumen sample collection.

4.7 References


33. **Chao A, Shen T-J.** 2010. Program SPADE (Species Prediction And Diversity Estimation).


4.8 Figures

Figure 4-1 A map of the full-length protozoal 18S rRNA gene, including variable (V1-V9) and rumen ciliate signature regions (SR1-SR4), and showing the respective amplicons of the three primer sets used in the present study.
Figure 4-2 Taxonomy and proportion of unique pyrosequences using NCBI (BLAST), by forward primers P-SSU-316F (Sylvester et al., 2004), GIC1080F (present study), and GIC1184F (present study).

All sequences used passed all quality assurance steps outlined in Methods.
Figure 4-3 Distribution of sequence identity percentages to known sequences based on BLAST by sequencing primer.
## 4.9 Tables

Table 4-1 Gastrointestinal tract ciliate protozoa primer specifications.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence (5'\rightarrow 3')</th>
<th>Tm* (50mM NaCl)</th>
<th>Dimer Formation</th>
<th>Hairpin Formation</th>
<th>End Stability</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIC758R</td>
<td>CAACTGTCTCTATKAAYCG</td>
<td>47.4°C</td>
<td>2 bp</td>
<td>2 bp</td>
<td>Medium</td>
</tr>
<tr>
<td>GIC1080F</td>
<td>GGGRAACTTACCAGGTCC</td>
<td>53.6°C</td>
<td>3 bp</td>
<td>3 bp</td>
<td>Medium</td>
</tr>
<tr>
<td>GIC1184F</td>
<td>TGTCTGGTTAATTCCGA</td>
<td>47.2°C</td>
<td>4 bp</td>
<td>3 bp</td>
<td>High</td>
</tr>
<tr>
<td>GIC1578R</td>
<td>GTGATRWGRTTTACTTTRT</td>
<td>42.1°C</td>
<td>2 bp</td>
<td>2 bp</td>
<td>High</td>
</tr>
</tbody>
</table>

*Melting temperature. K=G/T Y=C/T R=A/G W=A/T
Table 4-2 18S rRNA protozoal reference sequences used to determine genetic distance cutoff.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession #</th>
<th>Species</th>
<th>Accession #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alloiozona trizona</td>
<td>AB795026</td>
<td>Epidinium ecaudatum</td>
<td>AM158465</td>
</tr>
<tr>
<td>Amylovorax dehorityi</td>
<td>AF298817</td>
<td>Eremoplastron dilobum</td>
<td>AM158472</td>
</tr>
<tr>
<td>Amylovorax dogieli</td>
<td>AF298825</td>
<td>Eremoplastron neglectum</td>
<td>AM158473</td>
</tr>
<tr>
<td>Balantidium ctenophaerygodoni</td>
<td>GU480804</td>
<td>Eremoplastron rostratum</td>
<td>AM158469</td>
</tr>
<tr>
<td>Balantidium entozoon</td>
<td>EU581716</td>
<td>Eudiplodinium maggi</td>
<td>U57766</td>
</tr>
<tr>
<td>Bandia cribbi</td>
<td>AF298824</td>
<td>Gassovskiella galea</td>
<td>AB793783</td>
</tr>
<tr>
<td>Bandia deveneyi</td>
<td>AY380823</td>
<td>Helicozoster indicus</td>
<td>AB794981</td>
</tr>
<tr>
<td>Bandia smalesae</td>
<td>AF298822</td>
<td>Hemiprorodon gymnoprosthium</td>
<td>AB795028</td>
</tr>
<tr>
<td>Bandia tammar</td>
<td>AF298823</td>
<td>Isotricha intestinalis</td>
<td>U57770</td>
</tr>
<tr>
<td>Bitricha tasmaniensis</td>
<td>AF298821</td>
<td>Isotricha prostoma</td>
<td>AF029762</td>
</tr>
<tr>
<td>Blepharocerus hemiciliatus</td>
<td>AB795027</td>
<td>Latteuria media</td>
<td>AB794983</td>
</tr>
<tr>
<td>Blepharocorys angusta</td>
<td>AB794976</td>
<td>Latteuria polyastra</td>
<td>AB794982</td>
</tr>
<tr>
<td>Blepharocorys curvigula</td>
<td>AB534184</td>
<td>Macropodinium ennunensis</td>
<td>AF298820</td>
</tr>
<tr>
<td>Blepharocorys jubata</td>
<td>AB794977</td>
<td>Macropodinium yalabense</td>
<td>AF042486</td>
</tr>
<tr>
<td>Blepharocorys microcorys</td>
<td>AB794975</td>
<td>Neobalantidium coli</td>
<td>AF029763</td>
</tr>
<tr>
<td>Blepharocorys uncinata</td>
<td>AB530162</td>
<td>Ochterenatia appendiculata</td>
<td>AB794973</td>
</tr>
<tr>
<td>Bozasella sp</td>
<td>AB793744</td>
<td>Ophryoscolex caudatus</td>
<td>AM158467</td>
</tr>
<tr>
<td>Bundleia benbrooki</td>
<td>AB555711</td>
<td>Ophryoscolex purkynjei</td>
<td>U57768</td>
</tr>
<tr>
<td>Bundleia nana</td>
<td>AB555712</td>
<td>Ostracodinium clipeolum</td>
<td>AB536717</td>
</tr>
<tr>
<td>Bundleia postciliata</td>
<td>AB555709</td>
<td>Ostracodinium gracile</td>
<td>AB535662</td>
</tr>
<tr>
<td>Buxtonella sulcata</td>
<td>AB794979</td>
<td>Paraisotricha colpoidea</td>
<td>EF632075</td>
</tr>
<tr>
<td>Circodinium minimum</td>
<td>AB794974</td>
<td>Paraisotricha minuta</td>
<td>AB794984</td>
</tr>
<tr>
<td>Cochliatoxum perichtum</td>
<td>EF632078</td>
<td>Parentodinium sp.</td>
<td>AB530164</td>
</tr>
<tr>
<td>Cycloposthium bipalatum</td>
<td>AB530165</td>
<td>Polycosta roundi</td>
<td>AF298819</td>
</tr>
<tr>
<td>Cycloposthium edentatum</td>
<td>EF632077</td>
<td>Polycosta turniae</td>
<td>AF298818</td>
</tr>
<tr>
<td>Cycloposthium ishikawai</td>
<td>EF632076</td>
<td>Polidyinella mysorea</td>
<td>AB555710</td>
</tr>
<tr>
<td>Dasystricha ruminantium</td>
<td>U57769</td>
<td>Polyplastron multivesiculatum</td>
<td>U57767</td>
</tr>
<tr>
<td>Didesmis ovalis</td>
<td>AB795025</td>
<td>Prorodonopsis coli</td>
<td>AB795029</td>
</tr>
<tr>
<td>Diplopinum dentatum</td>
<td>U57764</td>
<td>Pseudoentodinium elephantis</td>
<td>AB794972</td>
</tr>
<tr>
<td>Diploplastron affine</td>
<td>AM158457</td>
<td>Raabena bella</td>
<td>AB534183</td>
</tr>
<tr>
<td>Ditoxum funinucleum</td>
<td>AB794091</td>
<td>Spirodinium equi</td>
<td>AB794092</td>
</tr>
<tr>
<td>Entodinium caudatum</td>
<td>U57765</td>
<td>Sulcoarcus pellucidulus</td>
<td>AB795024</td>
</tr>
<tr>
<td>Entodinium dubardi</td>
<td>AM158443</td>
<td>Tetraplastrum parvum</td>
<td>AB794969</td>
</tr>
<tr>
<td>Entodinium longinucleatum</td>
<td>AB481099</td>
<td>Triadinium caudatum</td>
<td>AB530163</td>
</tr>
<tr>
<td>Entodinium nanellum</td>
<td>JX876561</td>
<td>Tripalmaria dogieli</td>
<td>EF632074</td>
</tr>
<tr>
<td>Entodinium simplex</td>
<td>AM158466</td>
<td>Triplumaria selenica</td>
<td>AB533538</td>
</tr>
<tr>
<td>Epidinium caudatum</td>
<td>U57763</td>
<td>Troglodytella brassarti</td>
<td>AB437346</td>
</tr>
</tbody>
</table>
Table 4-3 Comparison of statistical parameters and results for three primer sets.

<table>
<thead>
<tr>
<th></th>
<th>P-SSU-316F GIC758R</th>
<th>GIC1080F GIC1578R</th>
<th>GIC1184F GIC1578R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sequences before quality assurance steps</td>
<td>19,886</td>
<td>10,143</td>
<td>8,611</td>
</tr>
<tr>
<td>Total sequences after quality assurance steps</td>
<td>12,326</td>
<td>6,070</td>
<td>8,265</td>
</tr>
<tr>
<td>Unique sequences used for BLAST</td>
<td>769</td>
<td>697</td>
<td>424</td>
</tr>
<tr>
<td>Trimmed sequence length (bases)</td>
<td>450 b</td>
<td>450 b</td>
<td>375 b</td>
</tr>
<tr>
<td>Subsampled unique sequences used for statistical analysis</td>
<td>424</td>
<td>424</td>
<td>424</td>
</tr>
<tr>
<td>Species-level distance within genera(^1)</td>
<td>0.036</td>
<td>0.039</td>
<td>0.042</td>
</tr>
<tr>
<td>Recommended % cutoff for species-level OTUs</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Observed species-level OTUs</td>
<td>48</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>CHAO</td>
<td>112</td>
<td>58</td>
<td>42</td>
</tr>
<tr>
<td>ACE</td>
<td>226</td>
<td>92</td>
<td>145</td>
</tr>
<tr>
<td>Shannon-Weaver</td>
<td>2.36</td>
<td>1.02</td>
<td>1.37</td>
</tr>
<tr>
<td>Good’s Coverage</td>
<td>0.94</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>Inverse Simpson</td>
<td>4.69</td>
<td>1.71</td>
<td>2.59</td>
</tr>
<tr>
<td>Genus-level distance across genera(^2)</td>
<td>0.087</td>
<td>0.079</td>
<td>0.096</td>
</tr>
<tr>
<td>Recommended % cutoff for genera-level OTUs</td>
<td>9%</td>
<td>8%</td>
<td>10%</td>
</tr>
<tr>
<td>Observed genus-level OTUs</td>
<td>15</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

\(^1\) Using near full-length 18S rRNA gene sequences from valid species, the species-level cutoff was calculated to be 0.031.

\(^2\) Using near full-length 18S rRNA gene sequences, the genus-level cutoff was calculated to be 0.071.

4.10 Supplemental Material

Supplemental Material from this publication includes four lower triangle charts of genetic distance comparisons, which are too large to be included in this text. Supplemental Material can be accessed from the journal Applied and Environmental Microbiology: http://aem.asm.org/content/early/2014/06/23/AEM.01644-14/suppl/DCSupplemental
CHAPTER 5  HIGH-THROUGHPUT DNA SEQUENCING OF THE MOOSE RUMEN FROM DIFFERENT GEOGRAPHICAL LOCATION REVEALS A CORE RUMINAL METHANOGENIC ARCHAEAAL DIVERSITY AND A DIFFERENTIAL CILIATE PROTOZOAL DIVERSITY.

Suzanne L. Ishaq¹,*, Monica A. Sundset², John Crouse³, and André-Denis G. Wright¹,⁴

¹ Department of Animal Science, University of Vermont, Burlington, Vermont, 05401
² Department of Arctic and Marine Biology, University of Tromsø, Tromsø, Norway, 9019
³ Alaska Department of Fish and Game, Soldotna, Alaska, 99669
⁴ Present address: School of Animal and Comparative Biomedical Sciences, University of Arizona, Tucson, Arizona 85721

*Contact: Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax) 1-802-656-8196.

Keywords: 16S rRNA/ 18S rRNA/ 454 Roche Titanium/ Alaska/ MiSeq/ Norway/ rumen/ Vermont
Moose rumen samples from Vermont, Alaska, and Norway were investigated for methanogenic archaeal and protozoal density using real-time PCR, and diversity using high-throughput sequencing of the 16S rRNA and 18S rRNA gene, respectively. Vermont moose showed the highest protozoal and methanogen densities. Alaskan samples had the highest percentages of *Methanobrevibacter smithii*, followed by the Norwegian samples. One Norwegian sample contained 43% *Methanobrevibacter thaueri*, while all other samples contained <10%. Vermont samples had large percentages of *Methanobrevibacter ruminantium*, as did two Norwegian samples. *Methanosphaera stadtmanae* represented one third of sequences in three samples. Samples were heterogeneous based on gender, geographic location, and weight class using AMOVA, but did not cluster significantly using PCoA or Unifrac. Two Alaskan moose contained >70% *Polyplastron multivesiculatum*, and one contained >75% *Entodinium* sp. Protozoa from Norwegian moose belonged predominantly (>50%) to the genus *Entodinium*, especially *Ent. caudatum*. Norwegian moose also contained a large proportion of sequences (25-97%) which could not be classified beyond Ophryoscolecidae. Protozoa from Vermont samples were predominantly *Eudiplodinium rostratum* (>75%), with up to 7% *Diploplastron affine*. Four of the eight Vermont samples also contained 5-12% *Entodinium* spp. Samples were heterogeneous based on AMOVA, PCoA, and Unifrac. Previously, Alaskan moose were speculated to consume a higher starch diet than Vermont moose, which were presumably consuming a higher forage diet.
5.2 Introduction

Previous investigations into the microorganisms in the rumen of the moose have focused on bacteria using cultivation [1] and high-throughput sequencing techniques [2, 3], or on protozoa using light microscopy [4–7] and high-throughput sequencing [8]. Methanogenic archaia in the rumen of moose have not been previously identified, nor have methanogens or protozoa from moose been compared across samples from different geographic locations. Methanogens and protozoa in the rumen are often found in intracellular or extracellular symbiotic associations involving hydrogen transfer from protozoa to methanogens. Previously, protozoa from the genera *Entodinium*, *Polyplastron*, *Epidinium*, and *Ophryoscolex* have been shown to interact with methanogens from the orders Methanobacteria and Methanococci [9].

The objectives of this research were to identify the methanogens and protozoa present in the rumen of moose from Alaska, Vermont, and Norway; to measure the density of methanogens and protozoa in these samples; to compare samples across geographic location, gender, and weight class to determine possible trends; and to compare samples to published studies on wild and domesticated ruminants. It was hypothesized that moose may have fewer total methanogens than domestic ruminants due to a fast rate of passage through the gastrointestinal tract [10]. In previous studies, age [3, 11, 12] and geographic location [3, 13] have played a role in differentiating core bacterial microbiomes, and it was also hypothesized that this would hold true for methanogens in the moose rumen.
However, reindeer in various geographic locations have been shown to have similar protozoal diversity, indicating that the host species may have been isolated long enough to develop a common profile regardless of geographic location of the host [14].

5.3 Methods

A total of 17 samples were collected from the rumen of moose in Vermont, USA (n=8) (Oct 2010), Troms County, Norway (n=6) (Sept-Oct 2011), and Soldotna, Alaska (n=3) (Aug 2012). Sample collection and DNA extraction were previously described [3]. Metadata for each sample collected, including gender, weight, approximate age, and coordinates of sample collection have also been published elsewhere [3]. Pooled samples from Alaska (n=3) were previously sequenced and described [8]. Samples were identified by location (Alaska=AK, Norway=NO, and Vermont=VT), host (m=moose), individual moose (1-8), and by sample material (r=rumen), consistent with previous publications [2, 3, 8].

PCR was performed on a C1000 ThermalCycler (Bio-Rad, Hercules, CA) using the Phusion kit (ThermoScientific, CA) to amplify rDNA. For methanogenic archaea, the V1 to V3 region of the archaeal 16S rRNA gene was amplified using primers 86F (5’-GCTCAGTAACACGTGG-3’) [15] and 471R (5’-GWRTTACCGCGGCKGCTG-3’) [16]. The protocol was as follows: initial denaturing at 98°C for 10 min, then 35 cycles of 98°C for 30s, 58°C for 30s, 72°C for 30s, then a final elongation step of 72°C for 6
min. For ciliate protozoa, the V3-V4 and signature regions 1-2 of the 18S rRNA gene were amplified using primers P-SSU-316F (5’-GCTTTCGWTGGTAGTGTATT-3’) [17] and GiC758R (5’-CAACTGTCTCTATKAAYCG-3’) [8] following previously described conditions [8]. PCR amplicons were verified on an agarose gel (100V, 60 min), and DNA bands were excised and purified as previously described [3]. Amplicons were sent to MR DNA Laboratories (Shallowater, TX) for MiSeq ver. 3 (methanogens) or Roche 454 pyrosequencing with Titanium (protozoa).

5.3.1 Sequence analysis

All sequences were analyzed using MOTHUR ver. 1.31 [18]. For methanogens, sequence analysis was as previously described [3], with the following modifications. Sequences were trimmed to a uniform length of 436 alignment characters (minimum 350 bases), and candidate sequences were aligned against the Ribosomal Database Project (RDP) reference alignment integrated into MOTHUR with the bacterial sequences removed. Sequences were classified using the k-nearest neighbor method against the full RDP alignment, which had been modified to include species-level taxonomy. A 2% genetic distance cutoff was used to designate species. For protozoa, sequence analysis was as previously described for primer set 1, using a 4% genetic distance cutoff to designate species [8]. For each sample, sequences were subsampled, and CHAO [19], ACE [20], Good’s Coverage [21] and the Shannon-Weiner Index [22] were calculated.
An analysis of molecular variance (AMOVA) and Unifrac [23] were used to compare the heterogeneity of samples.

5.3.2 Real-time PCR

Real-time PCR was used to calculate archael and protozoal densities in whole samples. DNA was amplified using a CFX96 Real-Time System (Bio-Rad, CA) and a C1000 Thermal Cycler (Bio-Rad, CA). Data were analyzed using CFX Manager Software ver. 1.6 (Bio-Rad, CA). The iQ SYBR Green Supermix kit (Bio-Rad, CA) was used: 12.5µL of mix, 2.5µl of each primer (40mM), 6.5µL of ddH₂O, and 1µL of the initial DNA extract diluted to approximately 10 ng/µL. For methanogens, the primers targeted the methyl coenzyme-M reductase A gene (mcrA), following the protocol by Denman et al. [24]. The internal standards for methanogens were a mix of Methanobrevibacter smithii, M. gottschalkii, M. ruminantium and M. millerae (R²=0.998).

For protozoa, the primers, PSSU316F and PSSU539R [17], targeted the 18S rRNA gene, following the protocol by Sylvester et al. [17], and the internal standards for protozoa were created in the laboratory using fresh rumen contents which were filtered through one layer of cheesecloth to remove large particles, and then the protozoa were allowed to separate for two hours at 39°C. Once a protozoal pellet was visible, 50 ml were drawn from the bottom of the funnel, and 1 volume of ethanol was added to fix the cells and DNA. The mix was centrifuged for 5 min at 2,000 x G, the pellet was washed with TE buffer (1MTris-HCl, 0.5 M EDTA, pH 8.0) and then centrifuged again. Cells were
counted microscopically using a Thoma Slide following the protocol by Dehority [5], 
($R^2=0.998$). Both protocols were followed by a melt curve, with a temperature increase 0.5°C every 10s from 65°C up to 95°C to check for contamination.

5.4 Results

5.4.1 Methanogens

A total of 141,368 sequences, of which, 47,370 were unique sequences, passed quality assurance steps. For each sample, between 22 and 330 OTUs were assigned using a 2% genetic distance cutoff, giving a total of 1,942 non-redundant OTUs. CHAO, ACE, Good’s Coverage and Shannon-Weaver Diversity for each sample are provided in Table 1. The Vermont samples showed the highest Shannon index, CHAO, and ACE, while the Norwegian samples showed the highest Good’s Coverage. The Alaskan samples showed the highest observed OTUs. Although there were few shared OTUs among samples, these shared OTUs represented a large number of shared sequences (Table 2). Comparing all 17 samples across different factors using AMOVA, groups were heterogeneous based on gender ($p<0.001$), geographic location ($p=<0.001$), and weight class ($p<0.001$). In contrast, samples did not cluster significantly based on gender or weight class using PCoA (Figure 1 A,C,E), although Vermont clustered separately from Norway and Alaska. When comparing samples using Unifrac, all samples again did not cluster significantly using either weighted (0.17, $p<0.001$) or unweighted (0.91, $p=0.20$)
parameters. However, 16 out of 136 pairwise sample comparisons were significantly different (p<0.001).

Vermont samples contained the highest mean density of methanogens at 1.3E+10, followed by Alaskan samples and Norwegian samples (5.19E+09 and 3.58E+09, respectively) (Table 3). Alaskan samples had the highest percentages of *Methanobrevibacter (Mbr.) smithii* (16 to 36%), followed by the Norwegian samples (10 to 24%) (Figure 2). The Norwegian sample NO1R, contained the highest percentage of *Mbr. thaueri* (43% of total sequences), while all other samples contained <10%. Vermont samples had large percentages of *Mbr. ruminantium* (27-51% of total sequences), as did the Norwegian samples NO3R and NO4R (40 and 41%, respectively) (Figure 2). *Methanosphaera stadmanae* was highest in NO5R (36%), VT8R (35%), and NO6R (34%) (Figure 2). Less than 36 sequences total were found of each of the following: *Methanocella, Methanospirillum, Methanolobus, Methanosarcina, Picrophilus, Methanobacterium, Mbr. curvatus, Mbr. cuticularis*, or Unclassified at the genus level (“Other”, Figure 2).

### 5.4.2 Protozoa

A total of 499,152 sequences, of which, 72,091 were unique sequences, passed quality assurance steps. For each sample, between 1 and 31 OTUs were estimated using a 4% genetic distance cutoff, giving a total of 110 non-redundant OTUs. CHAO, ACE, Good’s
Coverage and Shannon-Weaver Diversity index for each sample are provided in Table 1. Both Norwegian and Vermont samples had extremely high coverage (>0.97%), yet low Shannon diversity, CHAO and ACE values. Although there were few shared OTUs among samples, these shared OTUs represented a large number of shared sequences (Table 2). When comparing samples using Unifrac, samples clustered significantly using weighted (0.71, p<0.001) and unweighted (0.93, p<0.001) parameters. When comparing the Norway and Vermont samples across different factors using AMOVA, groups were heterogeneous based on gender (p<0.001), geographic location (p=<0.001), and weight class (p<0.001). This was also confirmed using PCoA for gender, location, and weight class (Figure 1B, D, F).

Vermont samples contained the highest mean density of protozoa at 4.70E+06, followed by Alaskan samples and Norwegian samples (3.83E+06 and 5.17E+04, respectively) (Table 3). Using a previously described reference alignment and taxonomy of valid protozoal sequences [8], protozoa were identified (Figure 3). Two Alaskan moose contained >70% Polyplastron multivesiculatum, and one contained >75% Entodinium sp. Protozoa from Norwegian moose belonged predominantly (>50% of total sequences) to the genus Entodinium, especially Ent. caudatum (Figure 3). Norwegian moose also contained a large proportion of sequences (25-97% of total sequences) which could not be classified beyond the Ophryoscolecidae family (Figure 3). Protozoa from Vermont samples were predominantly composed of Eudiplodinium rostratum (>75% of total...
sequences). Vermont samples also contained up to 7% *Diploplastron affine* (Figure 3). Many other species were identified in moose, with <1% each of the following identified: *Anoplodinium denticulatum*, *Dasytricha* spp., *Diplodinium dentatum*, *Enoploplastron triloricatum*, *Entodinium bursa*, *Ent. dubardi*, *Ent. furca dilobum*, *Ent. furca monolobum*, *Ent. longinucleatum*, *Ent. simplex*, *Epidinium caudatum*, *Epi. ecudatum caudatum*, *Epidinium* spp., *Eremoplastron dilobum*, *Eremoplastron rostratum*, *Eudiplodinium maggi*, *Isotricha intestinalis*, *Isotricha prostoma*, *Metadinium medium*, *Metadinium minorum*, *Ophryoscolex purkynjei*, *Ophryoscolex* spp., *Ostracodinium cipeolum*, *Ostracodinium dentatum*, *Ostracodinium gracile*, and *Ostracodinium* spp.

### 5.5 Discussion

The present study represents the first insight into the methanogenic archaeal diversity in the rumen of the moose. Two of three Alaskan moose, as well as two of six Norwegian moose had a larger proportion of methanogens belonging to the SGMT clade (*Mbr. smithii*, *Mbr. gottschalkii*, *Mbr. millerae*, and *Mbr. thaueri*). All eight Vermont moose, one Alaskan moose and four Norwegian moose had greater proportions of members of the RO clade (*Mbr. ruminantium* and *Mbr. olleyae*). Previously, the SGMT clade was show to be prevalent in alpaca [25], sheep [26], and reindeer [27]. As with bacteria in a previous study [3], Alaskan moose shared a large number of methanogenic sequences, followed by the Norwegian samples, and females shared more archaeal and protozoal
sequences than males. Unlike previously, the 202-300 kg weight class shared the greatest number of archaeal and protozoal sequences of all the weight classes.

*Methanobrevibacter smithii*, unlike many other methanogens, has been shown to grow at less than neutral pH [28], is often associated with high-calorie diets, has been shown to influence weight gain in rats [29], and has been shown to improve polysaccharide fermentation by bacteria [30, 31]. Conversely, *Mbr. ruminantium* has been associated with a low-energy (high forage) diet [32]. Previously, Alaskan moose were speculated to be on a higher starch/energy diet than those of Vermont moose, presumably on a higher forage/lower energy diet [2], which may account for the relatively high proportions of *Mbr. smithii* in Alaskan moose and high proportions of *Mbr. ruminantium* in Vermont moose in the present study. *Methanosphaera stadtmanae* has previously been associated with diets involving fruit, as they require methanol which is a byproduct of pectin fermentation, as has been previously seen in omnivores [33, 34] and ruminants [16, 35, 36]. Though distinct in terms of proportion of taxa present in each of the three moose populations, the samples were not statistically different between populations. This suggests that moose have a core methanogen microbiome, as has been suggested for protozoa in other host species [14].

In the present study, protozoa from Alaska, Norway, and Vermont were distinctly different. Norwegian samples were dominated by *Entodinium* spp., while Vermont
samples were dominated by *Polyplastron multivesiculatum* and *Eudiplodinium maggi*. The previously published data are limited by the fact that protozoal sequences were generated using mixed amplicons from three Alaskan moose, not sequenced individually [8].

Previously, using light microscopy, moose were shown to have primarily *Entodinium* spp., including *Ent. dubardi* and *Ent. longinucleatum* in Alaska [5], *Ent. dubardi* and other *Entodinium* spp. from Slovakia, and *Ent. dubardi* and *Epi. caudatum* in Finish Lapland [7]. More recently using high-throughput sequencing, moose in Alaska were shown to have a high percentage of *Polyplastron multivesiculatum*, and well as a variety of *Entodinium* and other species [8], which was also shown in the present study. The Norwegian samples had a high percentage of *Ent. caudatum*, *Ent. furca dilobum*, and other *Entodinium* species, giving them a similar profile to moose samples from Alaska [5], Finland [7], and Slovakia [6] using light microscopy. The Norwegian samples also contained a large proportion of sequences which could not be identified beyond the family level, indicating that these moose host novel ciliate species, or that no 18S rRNA sequences exist for previously identified species.

Given the markedly different protozoal populations found in Alaska, Vermont, and Norway, as well as the AMOVA analysis confirming statistically different groups, it may be concluded that moose do not have a typical protozoal diversity as do reindeer [14].
Factors such as diet [27, 37, 38], and weaning strategy [39] have an effect on numbers and type of protozoa. Previously, total protozoal counts were shown to be elevated in concentrate selectors [38], while *Entodinium* populations were decreased in animals fed a higher concentrate diet over those fed a roughage diet [38]. *Entodinium* spp. are a major source of starch digestion in the rumen, as well as bacterial digestion [40]. *Polyplastron multivesiculatum* produces xylanase and other carbohydrate-degrading enzymes [41], which allows it to break down hemicellulose in plant cell walls and contribute to fiber digestion. *Eudiplodinium* spp. also preferentially ingest structural carbohydrates [42, 43].

It has been shown that protozoal density affects methanogen density [37, 44], as the two microbial communities are often symbiotically associated with one another. In particular, *Polyplastron, Eudiplodinium maggi*, and *Entodinium caudatum* have been shown to have >40% association with methanogens [45]. More specifically, *Polyplastron* was recently shown to associate with *Methanosphaera stadtmanae* and *Mbr. ruminantium* [46].

Methanogen and protozoal densities in reindeer from Norway [47] averaged very closely to densities found in Norwegian moose, but lower than in Alaskan and Vermont moose. Domestic steers fed a roughage diet had an average density of 1.34E+09 for methanogens, which were predominantly *Methanobrevibacter* spp. [24], and which was lower than the present study. Holstein dairy cattle on a high-forage diet had an average density of 6.04E+05 for protozoa [17], which were 2 logs less than densities in moose. It
was also shown that densities decreased on a low-forage diet, and the dominant genus was *Entodinium* spp. [17]. Roughage diets in livestock have been shown to increase methane emissions [48], even when the roughage diets are not associated with altered methanogen densities [32, 49].

### 5.6 Acknowledgements

Acknowledgments The authors would like to acknowledge the Vermont Fish and Wildlife Department for sample collection logistics; Terry Clifford, Archie Foster, Lenny Gerardi, Ralph Loomis, Beth and John Mayer, and Rob Whitcomb for collection of Vermont moose samples; Dr. Even Jørgensen, University of Tromsø, and Dr. Helge K. Johnsen, University of Tromsø, for collection of Norwegian moose samples; and Dr. Kimberlee Beckmen of the Alaska Department of Fish & Game for collection of Alaskan moose samples.

### 5.7 References


6. Sládeček F: Ophryoscolecidae from the stomach of *Cervus elaphus L.*, *Dama dama L.*, and *Capreolus capreolus L.* *Vestn Csl Zool Spole* 1946, **10**:201–231.


5.8 Figures

Figure 5-1 PCoA of moose (A, C, and E) and protozoa (B, D, and F). A, B are colored for gender (red=female and blue=male) for methanogens and protozoa, respectively. C, D are colored for location (red=Alaska, green= Norway, and blue=Vermont) for methanogens and protozoa, respectively. E, F are colored for weight class (na= red, 1-100 kg= dark blue, 101-200= right facing green, 201-300= down facing green, 301-400= yellow, 400+= light blue) for methanogens and protozoa, respectively.
Figure 5-2 Moose rumen methanogen taxonomy of total sequences, per sample from Alaska, Norway and Vermont

“Other” represents the following taxa which has <1% Methanocella, Methanospirillum, Methanolobus, Methanosarcina, Picrophilus, Methanobacterium, Mbr. curvatus, Mbr. cuticularis, or Unclassified at the genus level.
Figure 5-3 Moose rumen protozoal taxonomy of total sequences, per sample from Norway and Vermont.

Other species constitute <1% each of the following: Anoplodinium denticulatum, Dasytricha spp., Diplodinium dentatum, Enoploplastron triloricatum, Entodinimum bursa, Ent. dubardi, Ent. furca dilobum, Ent. furca monolobum, Ent. longinucleatum, Ent. simplex, Epidinium caudatum, Epi. ecaudatum caudatum, Epidinium spp., Eremoplastron dilobum, Eremoplastron rostratum, Eudiplodinium maggi, Isotricha intestinalis, Isotricha prostoma, Metadinium medium, Metadinium minorum, Ophryoscolex purkynjei, Ophryoscolex spp., Ostracodinium clipeolum, Ostracodinium dentatum, Ostracodinium gracile, and Ostracodinium spp.
## 5.9 Tables

Table 5-1 Statistical measures per sample for methanogens (met) and protozoa (prot) in Alaska (AK), Norway (NO), and Vermont (VT).

Samples were subsampled using the smallest group for methanogens and protozoa. Species-level cutoff was 2% for methanogens and 4% for protozoa.

| Sample | Subsampled Sequences | | | | | |
|--------|----------------------|---|---|---|---|---|---|
|        | Total seqs | Total OTUs | OTUs | CHAO | ACE | Good's Coverage | Shannon-Weiner |
| **Methanogens** | | | | | | | |
| AKM1R  | 4366       | 152    | 18   | 161  | 14  | 0.93            | 0.47          |
| AKM2R  | 537        | 23     | 20   | 200  | 0   | 0.92            | 0.52          |
| AKM3R  | 53648      | 292    | 6    | 17   | 10  | 0.98            | 0.15          |
| Mean   | 19517      | 156    | 15   | 126  | 8   | 0.94            | 0.38          |
| NOM1R  | 506        | 22     | 22   | 232  | 0   | 0.91            | 0.56          |
| NOM2R  | 1830       | 79     | 19   | 163  | 5   | 0.98            | 0.14          |
| NOM3R  | 19990      | 70     | 6    | 16   | 5   | 0.94            | 0.39          |
| NOM4R  | 1355       | 33     | 15   | 115  | 0   | 0.94            | 0.39          |
| NOM5R  | 17130      | 293    | 12   | 56   | 39  | 0.96            | 0.30          |
| NOM6R  | 7106       | 70     | 9    | 32   | 41  | 0.97            | 0.24          |
| Mean   | 7986       | 95     | 14   | 102  | 19  | 0.95            | 0.35          |
| VTM1R  | 2351       | 102    | 26   | 262  | 328 | 0.90            | 0.70          |
| VTM2R  | 1803       | 63     | 23   | 213  | 105 | 0.91            | 0.59          |
| VTM3R  | 12855      | 99     | 11   | 46   | 111 | 0.96            | 0.31          |
| VTM4R  | 4477       | 149    | 22   | 219  | 47  | 0.91            | 0.56          |
| VTM5R  | 1180       | 82     | 38   | 470  | 675 | 0.85            | 1.02          |
| VTM6R  | 3142       | 155    | 29   | 325  | 275 | 0.89            | 0.77          |
| VTM7R  | 790        | 43     | 28   | 389  | 0   | 0.89            | 0.73          |
| VTM8R  | 8302       | 330    | 31   | 359  | 538 | 0.88            | 0.81          |
| Mean   | 4363       | 128    | 26   | 285  | 260 | 0.90            | 0.69          |
| **Protozoa** | | | | | | | |
| AKM1R  | 81387      | 31     | 1    | 1    | 0   | 1.00            | 0.01          |
| AKM2R  | 57698      | 31     | 1    | 1    | 0   | 1.00            | 0.01          |
| AKM3R  | 16200      | 12     | 1    | 1    | 0   | 1.00            | 0.01          |
| Mean   | 51762      | 25     | 1    | 1    | 0   | 1.00            | 0.01          |
| NOM1R  | 24605      | 1      | 1    | 1    | 0   | 1.00            | 0.00          |
| NOM2R  | 15468      | 4      | 2    | 2    | 0   | 0.99            | 0.05          |
| NOM3R  | 20351      | 9      | 3    | 4    | 0   | 0.97            | 0.14          |
| NOM4R  | 5354       | 1      | 1    | 1    | 0   | 1.00            | 0.00          |
| NOM5R  | 35189      | 7      | 2    | 2    | 0   | 0.99            | 0.06          |
| NOM6R  | 30145      | 7      | 2    | 2    | 0   | 0.99            | 0.08          |
| Samples Compared  | Methanogens |           |   |   |   | | Protozoa |           |   |   |   | |
|------------------|-------------|-----------|---|---|---| |             |-----------|---|---|---| |
| All samples (n=17)|             | 1         | 44967 | 1 | 48850 |
| AK samples (n=3)  |             | 2         | 19888 | 2 | 45572 |
| NO samples (n=6)  |             | 2         | 16227 | 2 | 1771 |
| VT samples (n=8)  |             | 2         | 11255 | 2 | 1635 |
| All Females (n=11)|             | 2         | 31300 | 2 | 47400 |
| All Males (n=6)   |             | 2         | 16070 | 2 | 1578 |
| 0-100 kg (NO1R, NO6R) |         | 2         | 2684  | 2 | 519 |
| 101-200 kg (NO2R, VT1R, VT3R) |   | 4         | 4804  | 2 | 528 |
| 201-300 kg (NO3R, NO5R, VT2R, VT6R) | | 2         | 14007  | 2 | 1384 |
| 301-400 kg (VT5R, VT7R, VT8R) | | 2         | 3729   | 2 | 541 |
Table 5.3 Real-time PCR results for methanogenic archaea and ciliate protozoa in Alaska (AK), Norway (NO), and Vermont (VT).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Corrected cells/ml rumen digesta archaea</th>
<th>Corrected cells/ml rumen digesta protozoa</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKM1R</td>
<td>3.33E+09</td>
<td>3.60E+06</td>
</tr>
<tr>
<td>AKM2R</td>
<td>1.91E+09</td>
<td>4.72E+05</td>
</tr>
<tr>
<td>AKM3R</td>
<td>1.03E+10</td>
<td>7.43E+06</td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>5.19E+09 (4.51E+09)</td>
<td>3.83E+06 (3.48E+06)</td>
</tr>
<tr>
<td>NO1R</td>
<td>8.66E+07</td>
<td>5.92E+03</td>
</tr>
<tr>
<td>NO2R</td>
<td>1.54E+08</td>
<td>1.10E+04</td>
</tr>
<tr>
<td>NO3R</td>
<td>1.95E+08</td>
<td>5.46E+04</td>
</tr>
<tr>
<td>NO4R</td>
<td>1.38E+08</td>
<td>7.26E+03</td>
</tr>
<tr>
<td>NO5R</td>
<td>2.17E+10</td>
<td>1.67E+05</td>
</tr>
<tr>
<td>NO6R</td>
<td>8.25E+08</td>
<td>6.45E+04</td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>3.58E+09 (8.76E+09)</td>
<td>5.17E+04 (6.20E+04)</td>
</tr>
<tr>
<td>VT1R</td>
<td>3.87E+09</td>
<td>2.14E+06</td>
</tr>
<tr>
<td>VT2R</td>
<td>1.88E+10</td>
<td>3.00E+06</td>
</tr>
<tr>
<td>VT3R</td>
<td>4.26E+10</td>
<td>9.02E+06</td>
</tr>
<tr>
<td>VT4R</td>
<td>1.98E+10</td>
<td>5.70E+06</td>
</tr>
<tr>
<td>VT5R</td>
<td>7.68E+09</td>
<td>6.53E+06</td>
</tr>
<tr>
<td>VT6R</td>
<td>8.93E+08</td>
<td>5.08E+05</td>
</tr>
<tr>
<td>VT7R</td>
<td>4.54E+09</td>
<td>5.50E+06</td>
</tr>
<tr>
<td>VT8R</td>
<td>6.02E+09</td>
<td>5.18E+06</td>
</tr>
<tr>
<td>Mean (SE)</td>
<td>1.3E+10 (1.38E+10)</td>
<td>4.70E+06 (2.70E+06)</td>
</tr>
<tr>
<td>Mean all (SE)</td>
<td>7.36E+09 (4.95E+09)</td>
<td>2.86E+06 (2.47E+06)</td>
</tr>
</tbody>
</table>
CHAPTER 6  DESCRIPTION OF STREPTOCOCCUS ALCIS, SP. NOV., AND STREPTOCOCCUS VERMONTENSIS, SP. NOV., AND PROPOSAL OF THREE NEW SUBSPECIES OF STREPTOCOCCUS GALLOLYTICUS ISOLATED FROM THE RUMEN OF MOOSE (ALCES ALCES) IN VERMONT.

Suzanne L. Ishaq1*, Doug G. Reis2, Hannah M. Lachance1, and André-Denis G. Wright1,3

1 Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, 570 Main St., Burlington, Vermont 05405

2 Department of Microbiology and Molecular Genetics, College of Agriculture and Life Sciences, University of Vermont, 95 Carrigan Drive., Burlington, Vermont 05405

3 Present Address: School of Animal and Comparative Biomedical Sciences, College of Agriculture and Life Sciences, University of Arizona, 1117 E. Lowell Street, Tucson, Arizona. 85721

* Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax) 1-802-656-8196.

Email addresses:  SLI: slpelleg@uvm.edu, DR: dreis@uvm.edu, HL: hannah.lachance@uvm.edu, ADGW: adwright@email.arizona.edu

Contents Category: New Taxa (Firmicutes)

Keywords/Cross-reference: 16S rRNA gene, anaerobic culturing, bacteria

GenBank ASCN: KP009806-KP009843
6.1 Summary

Standard anaerobic and aerobic culturing techniques were used to isolate, characterize, and investigate the functional abilities of 37 isolates of *Streptococcus* bacteria from the rumen of the North American moose. Isolates were able to grow at higher temperatures and salinities than previously described cultivars, and at a wider range of pH. All 37 isolates produced acid from fructose, galactose, glucose, glycerol, lactose, maltose, mannitol, and sucrose. Variable numbers of isolates produced acid from N-acetylglucosamine, arabinose, cellobiose, cellulose, inulin, mannose, melibiose, raffinose, and salicin. Twenty-nine isolates produced a ropy exopolysaccharide, three reduced tellurite, and one isolate produced indole from tryptophan. Two isolates did not produce biogenic amines from amino acids. Twenty-four isolates are new strains of *S. gallolyticus* subsp. *gallolyticus*, and can produce acid from glycogen, inulin, mannitol, melibiose, and raffinose. Three isolates are new strains of *S. gallolyticus* subsp. *macedonicus*, which are unable to produce acid from glycogen, inulin, mannitol, melibiose, and raffinose. Nine isolates were not able to produce acid from glycogen or raffinose. However, some were able to produce acid from inulin, mannitol or melibiose. Based on the differential biochemical capabilities, DNA-DNA hybridizations, seven house-keeping genes, and genetic distances, based upon 16S rRNA gene sequence, of 10 isolates, two new species: *S. alcis* sp. nov., and *S. vermontensis* sp. nov.; and three new subspecies of *S. gallolyticus* are proposed: *S. gallolyticus* subsp. *mannosilyticus* subsp.
S. gallolyticus subsp. melibiosilyticus subsp. nov., and S. gallolyticus subsp. ruminantium subsp. nov.

This study used classical culturing techniques to isolate and characterize isolates of Streptococcus gallolyticus from the rumen of the North American moose (Alces alces). It was hypothesized that isolates from the moose rumen would be functionally distinct from previously identified bacteria, and that they might be appropriate for food production. Only a few studies have been published on the bacteria in the ruminal environment of moose (Ishaq & Wright, 2014, 2012; Dehority, 1986).

Several species of the lactic acid bacteria (LAB) Streptococcus are non-pathogenic and have been extremely important economically in the production of dairy food products due to the secretion of an exopolysaccharide (EPS) that gives dairy byproducts desired textures (Bolotin et al., 2004; Georgalaki et al., 2000; McSweeney, 2004; Papadimitriou et al., 2012; Tsakalidou et al., 1998; Vincent et al., 2001). Streptococcus gallolyticus can tolerate high-salinity environments, like those found in cheese making (Beresford et al., 2001). Many species of Streptococcus possess decarboxylase enzymes, which make them capable of producing biogenic amines (i.e. biologically active molecules) from different precursor amino acids. Biogenic amines can cause a variety of adverse gastrointestinal and cardiovascular symptoms depending on the amine and the dose. Streptococcus gallolyticus isolates which produced one or
more of these amines would be considered unfit candidates for food production applications (Ladero et al., 2010; McSweeney, 2004).

Isolates were cultured from whole rumen digesta samples, which were collected fresh, during the October 2010 hunting season in Vermont, with permission of licensed hunters through the Vermont Department of Fish and Wildlife. For information on the moose age and weight, see Ishaq and Wright (2012). All isolations were performed on M8 agar (Bryant & Robinson, 1961; Dehority & Grubb, 1976) plates inside an anaerobic chamber (90% nitrogen, 5% hydrogen, 5% carbon dioxide) (COY Laboratories, Michigan, US). Samples were serially diluted, plated with 5 replicates, and monitored for up to 1 week. Colonies were picked and re-isolated on fresh media until pure using gram staining and colony morphology measurements, and monocultures were identified using near full-length 16S rRNA gene sequencing performed at the University of Vermont Cancer Center DNA Analysis Facility (Burlington, Vermont, US). The bacterial 16S rRNA gene was amplified using the universal bacterial primers 27F and 1494R (Lane, 1991).

PCR was performed using the iTaq DNA Polymerase kit (Bio-Rad, California, US) per kit instructions. PCR conditions were: initial denaturation of 94°C for 5 min, then 33 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min, followed by a final extension of 72°C for 6 min. PCR was performed on a C1000 thermal cycler (Bio-Rad, California, US). Recalcitrant isolates were first extracted using the DNA extraction protocol in the QIAamp DNA Stool Mini Kit (QIAGEN, Maryland, US).
Amplification was verified by 1% agarose gel electrophoresis (100V for 60 min), and the remaining PCR product was enzymatically cleaned with ExoSAP-IT (Affymetrix, California, US). Cycle sequencing primers used included 27F, 1492R, 1494R, 907R (Lane, 1991), and 907F (Blackall et al., 1995). Sequences were proofread using ChromasPro ver. 1.7.5 (Technelysium Pty. Ltd., Australia), and aligned using the CLUSTALW algorithm in MEGA ver. 5.05 (Tamura et al., 2011). The alignment was refined by eye, and then used to calculate pairwise genetic distance using the Kimura 2-parameter model (Kimura, 1980). Sequences were identified using GenBank’s Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990), and compared to published sequences of lactic acid bacteria from NCBI using a neighbor-joining tree generated with MEGA (Figure 1).

Near full-length sequences were generated for each isolate and deposited in NCBI under accession numbers KP009806-KP009843. Sequences had 99% identity to known sequences of S. gallolyticus subsp. macedonicus using BLAST. When clustered using a neighbor-joining tree, 37 isolates clustered along with S. gallolyticus subsp. macedonicus\(^{(T)}\) (Figure 1). Mean genetic identity among the isolates was 99.7% (range 97.7–100%). The mean sequence identity between individual isolates and S. gallolyticus subsp. macedonicus\(^{(T)}\) was 99.5% (range 98.4–99.7%), while the mean sequence identity between individual isolates and S. gallolyticus subsp. gallolyticus\(^{(T)}\) was 99% (range 97.6–99.27%). Thirty-six of 37 isolates had a higher percent identity to S. gallolyticus subsp. macedonicus\(^{(T)}\) than to S. gallolyticus subsp. gallolyticus\(^{(T)}\), the exception being
The novel isolates had the following sequence identity to *S. gallolyticus* subsp. *gallolyticus* (VTM3R19<sup>T</sup>, 99%; VTM4R28<sup>T</sup>, 99.2%; VT1R31<sup>T</sup>, 95.5%; VTM3R37<sup>T</sup>, 94.8%; VTM1R54<sup>T</sup>, 95.4%) and *S. gallolyticus* subsp. *macedonicus* (VTM3R19<sup>T</sup>, 99.3%; VTM4R28<sup>T</sup>, 99.4%; VT1R31<sup>T</sup>, 99.2; VTM3R37<sup>T</sup>, 98.5%; VTM1R54<sup>T</sup>, 98.8%).

Five isolates which were proposed new taxa were also classified based on seven housekeeping genes as previously described [204]: guanylate kinase, gmk (AB829357-AB829373); peroxide resistance, dpr (AB829337-AB829356); DNA topoisomerase IV subunit A, parC (AB829398-AB829412); phosphate acetyltransferase, pta (AB829413-AB829429); dihydrotase, pyrC (AB829430-AB829444); DNA repair protein, recN (AB829451-AB829472); and RNA polymerase sigma factor, rpoD (AB829374-AB829397). Isolates clustered separately from reference sequences for gmk, parC, pta, recN (Figures 2A, 2C, 2D, 2F), and three of five clustered separately for rpoD (Figure 2G).

Monocultures were maintained on M8 media plates with sodium azide to prevent potential contamination growth of gram negative species, then transferred to MRS media (Downes & Ito, 2001; De Mann *et al.*, 1960) for the duration of testing. Before each test, isolates were subcultured in MRS broth (1% v/v inoculation) for 24 h, and tests were run in triplicate. Stock aliquots of isolates were mixed with 80% glycerol and stored at -80°C. Isolates were given unique identifiers (i.e. VTM3R11) containing the following abbreviations: Vermont (VT), moose (M), individual number (1–4), and
rumen (R), as well as isolate number. All isolate colonies showed similar morphology: small, white, irregular to round colonies with an opaque, glistening and butyrous appearance. Cell morphologies were consistently non-motile, gram-positive cocci in small chains or groups. All isolates were catalase negative.

Optimal growth parameters were determined by incubating isolates for 24 h at various temperatures (25–49°C), pH (5.0–10.0) (adjusted prior to autoclaving), or salinities (0–9% NaCl). Optical density (absorbance, 600 nm) was used to determine relative growth using a Spectronic 200 (ThermoScientific, CA) (Georgalaki et al., 2002). Optimal growth ranges were set as isolates measuring >0.5% absorbance. Optimal temperature ranged from 31°C to 39°C, optimal salinity ranged from 0 to 3% NaCl, and optimal pH ranged from pH 5.7 to 7.7 (Supplemental Figure 1). Six isolates were able to grow above 1% absorbance at 25°C (VTM3R15, VTM1R33, VTM3R37<sup>T</sup>, VTM4R46, VTM4R49, VTM4R51), and seven isolates were able to grow above 0.5% at 49°C (VTM3R11, VTM3R26, VTM1R44, VTM4R46, VTM4R49, VTM1R50, VTM4R54). Six isolates were able to grow above 0.5% at pH 5.0 and above 1% absorbance at pH 10.0 (VTM3R15, VTM2R16, VTM3R37<sup>T</sup>, VTM3R40, VTM4R46). Only three isolates were able to grow above 0.5% at 9% NaCl (VTM1R54<sup>T</sup>, VTM3R42, VTM4R13). In the present study, isolates tolerated high salinity environments, very high and low pH ranges, and high and low temperatures: higher than previously described <i>S. macedonicus</i> (sic) isolates (Tsakalidou et al., 1998).
Heat tolerance was tested by incubating 48 h old cultures in a 60°C water bath for 30 min, then inoculating onto MRS agar plates and incubating at 37°C for up to 72 h to observe for growth. Six isolates showed no change in growth after heat shock (VTM1R14, VTM3R15, VTM4R46, VTM4R49, VTM4R51, VTM4R54), and three isolates grew minimally (VTM1R44, VTM1R48, VTM3R32).

Ruthenium red (RR) plates (Stingele & Mollet, 1995) were used to determine if isolates produced a ropy exopolysaccharide, which prevents RR staining the cell wall of isolates. Because RR was ineffective at differentiating isolates if added to the media before autoclaving as specified previously (Stingele & Mollet, 1995), an additional 0.2 g/L RR was added after autoclaving (i.e. 0.28 g/L). Twenty-nine isolates exhibited a ropy exopolysaccharide, which is preferable in dairy byproducts (Table 1).

Isolates which produced a ropy exopolysaccharide then had extracellular protein quantified using skimmed milk medium (SMM) (Dabour & LaPointe, 2005; De Vuyst et al., 1998). The combined protocol is as follows: cultures were heat treated at 90°C for 15 min to inactivate enzymes, then centrifuged at 12,000 x g for 20 min at 4°C to pellet bacterial cells, and 1 volume of 20% trichloroacetic acid was added to the supernatant to precipitate protein. The supernatant was centrifuged at 16,000 x g for 20 min at 4°C to pellet proteins. The supernatant was removed, the pellet air-dried, and then suspended in ddH$_2$O for the protein assay. The protein assay was performed using the Bio-Rad Protein Assay Kit, (Bio-Rad, California, US), following the manufacturer’s instructions. Casein
was used to make protein standards, and assays were read using a Spectronic 200 (ThermoScientific, California, US).

Thirteen isolates produced between 0.5–1 mg/ml of protein each (Table 1). Data are visualized in Supplemental Figure 2. The quality and quantity of EPS produced by cultures are important for a successful consumer food product, as it contributes to product structure, mouth feel, and texture (Folkenberg et al., 2005), especially in low-fat dairy products, as it maintains the texture and structure that would normally be provided by fat, and can be used in applications where stabilizers are unavailable for use (Patel & Prajapati, 2013). In previous studies, some isolates produced no EPS (Georgalaki et al., 2000).

To measure acid production, isolates were subcultured in skim milk (10% w/v), and acid production was recorded as the pH value at 1 h, 6 h and 24 h (Georgalaki et al., 2000), (Table 2), and is visualized in Supplemental Figure 3. In the present study, isolates produced a similar amount of acid after 6 h of incubation as previously described isolates (Georgalaki et al., 2000). Isolates were subcultured onto MRS agar plates containing the pH indicator chlorophenol red, to test for acidic byproducts from different carbohydrates (Georgalaki et al., 2000). All isolates produced acidic byproducts from D-fructose, galactose, D+-glucose, glycerol, lactose, maltose and sucrose, and a variable number produced acids from D-mannose, D-arabinose, salicin, N-acetylglucosamine, cellobiose, melibiose, inulin, D-raffinose, and cellulose (Table 2). Streptococcus gallolyticus has previously been shown to utilize cellobiose (Chamkha et al., 2002), and a presumptive
gene resembling that of a *Ruminococcus albus* endoglucanase has previously been identified in *S. gallolyticus* (Rusniok *et al.*, 2010). *Streptococcus gallolyticus* subsp. *gallolyticus* can produce acid from glycogen, inulin, mannitol, melibiose, pullulan and raffinose, all of which previous strains of *S. gallolyticus* subsp. *macedonicus* were unable to do (Osawa *et al.*, 1995; Schlegel, 2003; Tsakalidou *et al.*, 1998).

Isolates were tested on mannitol media for their ability to metabolize mannitol and tolerate potassium tellurite. All isolates fermented mannitol, and two reduced tellurite and produced black precipitate (VTM3R11, VTM3R17). One isolate did not tolerate potassium tellurite and was unable to grow (VTM1R33). Isolates were grown for 14 d to test for the production of indole from tryptophan using Kovac’s reagent (Kovacs, 1928) added to the culture broth. No isolates produced indole from tryptophan.

To determine whether isolates could hydrolyze esculin (and produce a black precipitate), strains were plated on esculin media with and without the presence of bile salts. Three isolates could not tolerate bile salts and exhibited no growth (VTM1R27, VTM1R33, VTM1R50), while 35 isolates tolerated bile salts and hydrolyzed esculin. Without bile salts, all 37 isolates were capable of hydrolyzing esculin. *Streptococcus gallolyticus* subsp. *gallolyticus* can hydrolyze esculin (Osawa *et al.*, 1995; Schlegel, 2003; Tsakalidou *et al.*, 1998).

Isolates were subcultured into Simmon’s Citrate slants (Simmons, 1926) for 7d, and observed for bacterial growth and color change, to indicate the ability use citrate as a carbon source and ammonia as a nitrogen source (Georgalaki *et al.*, 2000). Fifteen
isolates used citrate as their sole source of carbon and ammonium ions as their source of nitrogen (Table 1). To test ability to reduce nitrate to nitrite, isolates were subcultured into nitrate broth for 3 and 7 d and then tested for color change using potassium iodine strips moistened with 1N HCl. Three isolates showed variable ability to produce nitrite from nitrate after 3 d, and six more (n=12) isolates were positive after 7 d. Twenty-five isolates showed lipase activity (Table 1), as determined by halo formation on MRS agar plates (pH 6.8) with tributyrin (1% v/v) and arabic gum (1% w/v) (Georgalaki et al., 2000).

To test biogenic amine production, isolates were inoculated on media containing a precursor amino acid (2% w/v, ornithine, histidine, lysine, tyrosine) (Joosten & Northolt, 1989), and observed for blue/green color formation around colonies (Georgalaki et al., 2000). Most isolates produced biogenic amines from some precursor amino acids, 11 isolates produced all four amines, and two produced none (Table 1). Biogenic amines may be used to create functional foods, as some are involved in immune response, cell growth, and homeostasis regulation (Ladero et al., 2010). However, certain biogenic amines can be toxic in high concentrations and their presence in certain foods is not preferred (Ladero et al., 2010). In the present study, isolates created biogenic amines from histidine, lysine, ornithine, and tyrosine. In a previous study, strains of Streptococcus produced little to no biogenic amine (Georgalaki et al., 2000).

Using biochemical profiles, as well as 16S sequencing, the present study classified 24 isolates as S. gallolyticus subsp. gallolyticus and 3 isolates as S. gallolyticus
subsp. macedonicus. Based on the differential biochemical capabilities, DNA-DNA hybridizations, seven house-keeping genes, and 16S rRNA genetic distances of 10 isolates, two new species, *S. alcis* sp. nov. and *S. vermontensis* sp. nov., and three new subspecies of *S. galloyticus* are proposed: *S. galloyticus* subsp. *mannosilyticus* subsp. nov., *S. galloyticus* subsp. *melibiosilyticus* subsp. nov., and *S. galloyticus* subsp. *ruminantium* subsp. nov.

For the novel isolates, the type strain was tested for hemolysis pattern using 5% sheep’s blood on tryptic soy agar (TSA) plates (ThermoScientific, California, US). Of the two proposed novel species and three proposed novel subspecies, four were alpha-hemolytic, and *S. galloyticus* subsp. *melibiosilyticus* subsp. nov. (VTM3R37T) was non-hemolytic. Each type strain was also genotypically characterized by DNA-DNA hybridization. G+C content and DNA-DNA hybridization were calculated using a C1000 thermal cycler with CFX96 real-time system (Bio-Rad, California, US) and previously published protocols (Bowman *et al*., 1998; Moreira *et al*., 2011). The reference strains *Streptococcus galloyticus* subsp. *galloyticus* (T) (ATCC 700065) (Osawa *et al*., 1995) and *S. galloyticus* subsp. *macedonicus* (T) (ATCC BAA-249) (Tsakalidou *et al*., 1998) were used. DNA-DNA hybridization was calculated using the change in melting temperature (ΔTm) between the reference and hybrid strains (Moreira *et al*., 2011) using regression equations created with previously published DNA-DNA hybridization data (Schlegel *at al*., 2003). DDH, GC content, and ΔTm are presented in Table 3. G+C content ranged from 37.7 to 38.1%, ΔTm was between 0.3 and 2.6°C, and % DNA-DNA hybridization
to reference strains was between 73 and 91%. The species-level cutoff for DNA-DNA hybridization is 70%. However, previously published data on DNA-DNA hybridization ranged from 50–100% for strains of *S. gallolyticus* subsp. *gallolyticus*, and 54–100% for strains of *S. gallolyticus* subsp. *macedonicus* (Schlegel et al. 2003), indicating a high degree of genetic variability even within a subspecies.

### 6.2 Description of *Streptococcus alcis* sp. nov.

*Streptococcus alcis* (al’cis. N.L. gen. n. *alcis* named after the moose, *Alces alces*, the source of the type strain). Cells are Gram-positive cocci, occurring in pairs or short chains, non-motile, non-sporulating, and catalase-negative. Colonies are circular, 1 mm in diameter after 24 h at 37 °C, and white to unpigmented. Growth is enhanced in a 5% CO₂ atmosphere, and occurs in MRS broth without gas production. No growth in 6±0.5% (w/v) NaCl broth. Alpha hemolytic on 5% blood agar. The type strain of *S. alcis* VTM1R28^T (= ATCC-00408-01^T = DSM XXXX^T =NCBI KP009833) was isolated from the rumen of a 1-year old female moose in Vermont, USA. Produces acid from N-Acetylglucosamine, arabinose, cellobiose, cellulose, fructose, galactose, glucose, glycerol, glycogen, inulin, lactose, maltose, mannose, melibiose, raffinose, salicin and sucrose. Exopolysaccharide is ropy in consistency. Biogenic amines produced from histidine and lysine. Lipase is produced. It is tolerant of bile salts and can hydrolyze esculin. It can use citrate as a sole carbon source and ammonium ions as a sole nitrogen source.
source. Characteristics useful in their differentiation from related organisms and also in the delineation between the two subspecies are listed in Tables 1–3.

### 6.3 Description of Streptococcus vermontensis sp. nov.

*Streptococcus vermontensis* (ver.mont.en’sis. N.L. masc. adj. *vermontensis* named after the U.S. state where the moose were captured, the source of the type strain). Cells are Gram-positive cocci, occurring in pairs or short chains, non-motile, non-sporulating, and catalase-negative. Colonies are circular, 1 mm in diameter after 24 h at 37 °C, and white to unpigmented. Growth is enhanced in a 5% CO₂ atmosphere, and occurs in MRS broth without gas production. No growth in 6±0.5% (w/v) NaCl broth. Alpha hemolytic on 5% blood agar. The type strain of *S. vermontensis* VTM3R19ᵀ (= ATCC-00408-06ᵀ = DSM XXXXXᵀ =NCBI KP009835) was isolated from the rumen of a 2-year old male moose in Vermont, USA. Produces acid from N-acetylglucosamine, arabinose, cellobiose, fructose, galactose, glucose, glycerol, lactose, maltose, mannose, melibiose, salicin and sucrose. Exopolysaccharide is non-ropy in consistency. Biogenic amines produced from histidine, lysine, and ornithine. It is tolerant of bile salts and can hydrolyze esculin. Characteristics useful in their differentiation from related organisms and also in the delineation between the two subspecies are listed in Tables 1–3.
6.4 Description of Streptococcus gallolyticus subsp. mannosilyticus subsp. nov.

*Streptococcus gallolyticus* subsp. *mannosilyticus* (man.no.si.ly'ti.cus. N.L. neut. n. mannosum mannose; N.L. masc. adj. lyticus (from Gr. neut. adj. lutikos), able to loosen, able to dissolve; N.L. masc. adj. mannosilyticus breaking down mannose). Cells are Gram-positive cocci, occurring in pairs or short chains, non-motile, non-sporulating, and catalase-negative. Colonies are circular, 1 mm in diameter after 24 h at 37 °C, and white to unpigmented. Growth is enhanced in a 5% CO₂ atmosphere, and occurs in MRS broth without gas production. Some growth in 6±0.5% (w/v) NaCl broth. Alpha hemolytic on 5% blood agar. The type strain of *S. gallolyticus* subsp. *mannosilyticus* VTM1R31ᵀ (=ATCC-00408-03ᵀ = DSM XXXXXᵀ = NCBI KP009837) was isolated from the rumen of a 1-year old female moose in Vermont, USA. Produces acids from fructose, galactose, glucose, glycerol, lactose, maltose, mannose, and sucrose. Acid production from N-Acetylglucosamine, arabinose, cellobiose, cellulose, and salicin is variable. Exopolysaccharide is ropy or non-ropy in consistency. Biogenic amines produced from lysine, and variably from histidine, ornithine, and tyrosine. Lipase is variably produced. It is tolerant of bile salts and can hydrolyze esculin. Ability to use citrate as a sole carbon source and ammonium ions as a sole nitrogen source is variable. Three strains of this subspecies were also isolated from the moose rumen: VTM3R15, VTM1R44, and VTM4R49. Characteristics useful in their differentiation from related organisms and also in the delineation between the two subspecies are listed in Tables 1–3.
6.5 Description of Streptococcus gallolyticus subsp. melibiosilyticus subsp. nov.

*Streptococcus gallolyticus* subsp. *melibiosilyticus* (me.li.o.si.ly'ti.cus. N.L. neut. n. melibiosum melibiose; N.L. masc. adj. lyticus (from Gr. neut. adj. lutikos), able to loosen, able to dissolve; N.L. masc. adj. melibiosilyticus breaking down melibiose). Cells are Gram-positive cocci, occurring in pairs or short chains, non-motile, non-sporulating, and catalase-negative. Colonies are circular, 1 mm in diameter after 24 h at 37 °C, and white to unpigmented. Growth is enhanced in a 5% CO₂ atmosphere, and occurs in MRS broth without gas production. Some growth in 6±0.5% (w/v) NaCl broth. Non (gamma) hemolytic on 5% blood agar. The type strain of *S. gallolyticus* subsp. *melibiosilyticus* VTM3R37T (=ATCC-00408-04T = DSM XXXXXT =NCBI KP009841) was isolated from the rumen of a 1-year old female moose in Vermont, USA.Produces acid from arabinose, cellobiose, cellulose, fructose, galactose, glucose, glycerol, lactose, maltose, mannose, melibiose, salicin and sucrose. Acid production from N-Acetylglucosamine is variable. Exopolysaccharide is ropy in consistency. Biogenic amines produced from lysine, and variably from histidine, ornithine, and tyrosine. Lipase is variably produced. It is tolerant of bile salts and can hydrolyze esculin. Ability to use citrate as a sole carbon source and ammonium ions as a sole nitrogen source is variable. One strain of this subspecies was also isolated from the moose rumen, VTM3R23. Characteristics useful in their differentiation from related organisms and also in the delineation between the two subspecies are listed in Tables 1–3.
6.6 Description of *Streptococcus gallolyticus* subsp. *ruminantium* subsp. *nov.*

*Streptococcus gallolyticus* subsp. *ruminantium* (ru.mi.nan’ti.um. L. part adj. ruminans -antis, ruminating; N.L. pl. gen. n. ruminantium, of ruminants) named after the ruminant rumen/forestomach where the bacteria resided. Cells are Gram-positive cocci, occurring in pairs or short chains, non-motile, non-sporulating, and catalase-negative. Colonies are circular, 1 mm in diameter after 24 h at 37 °C, and white to unpigmented. Growth is enhanced in a 5% CO₂ atmosphere, and occurs in MRS broth without gas production. Some growth in 6±0.5% (w/v) NaCl broth. Alpha hemolytic on 5% blood agar. The type strain of *S. gallolyticus* subsp. *ruminantium* VTM1R54^T^ (=ATCC-00408-05^T^ = DSM XXXXX^T^ =NCBI KP009842) was isolated from the rumen of a 1-year old female moose in Vermont, USA. Produces acid from arabinose, fructose, galactose, glucose, glycerol, lactose, maltose, and sucrose. Acid production from cellulose and melibiose is variable. Exopolysaccharide is ropy or non-ropy in consistency. Biogenic amines produced from lysine, ornithine, and tyrosine. Lipase is produced. It is tolerant of bile salts and can hydrolyze esculin. One strain (VTM4R54) of this subspecies was also isolated from the rumen of moose in Vermont, US. Characteristics useful in their differentiation from related organisms and also in the delineation between the two subspecies are listed in Tables 1–3.

6.7 Conflict of interest

The authors declare no conflict of interest.

179
6.8 Acknowledgements

The authors would like to thank Sam Rosenbaum, Ken Wesley, and Emma Hurley for their help in preparing culture media, collecting optical density data, and for general lab maintenance; the Vermont Fish and Wildlife Department for their help in rumen sample collection logistics; and Terry Clifford, Archie Foster, Lenny Gerardi, Ralph Loomis, Beth and John Mayer, and Rob Whitcomb for collection of rumen samples.

6.9 References


6.10. Figures

Figure 6-1 Neighbor-joining tree comparing isolates to type isolates to lactic acid bacteria. Tree was generated using MEGA ver. 5.05 (Tamura et al., 2011) and the Kimura 2-parameter model (Kimura, 1980). (T) = type strain, and numbers at the nodes represent bootstrap values. *Lactobacillus acidophilus* (AB680529) and *Enterococcus faecium* (AJ301830) were used as out-groups.
Figure 6-2 UPGMA trees comparing novel isolates using the housekeeping genes gmk (A), dpr (B), parC (C), pta (D), pyrC (E), recN(F), and rpoD (G).
Figure 6-2 UPGMA trees comparing novel isolates using the housekeeping genes gmk (A), dpr (B), parC (C), pta (D), pyrC (E), recN(F), and rpoD (G).
Figure 6-2 UPGMA trees comparing novel isolates using the housekeeping genes gmk (A), dpr (B), parC (C), pta (D), pyrC (E), recN(F), and rpoD (G).
Figure 6-2 UPGMA trees comparing novel isolates using the housekeeping genes gmk (A), dpr (B), parC (C), pta (D), pyrC (E), recN(F), and rpoD (G).
6.11 Tables

Table 6-1 Ability of isolates to produce biogenic amines from selected amino acids, produce lipase, metabolize citrate, produce a ropy or non-ropy exopolysaccharide, and extracellular protein production.

For biogenic amine production from ornithine, histidine, tyrosine and lysine, and lipase production results are presented as positive (+), slight positive (~), and negative (-). Simmon’s Citrate results are presented as positive (+) or negative (-) for both butt (anaerobic) and slant (aerobic) portions, as well as color change (B). Exopolysaccharide results are presented by type: ropy or non-ropy (NR).

<table>
<thead>
<tr>
<th>Isolates</th>
<th>Ornithine</th>
<th>Histidine</th>
<th>Tyrosine</th>
<th>Lysine</th>
<th>Lipase</th>
<th>Simmons (b/s)</th>
<th>EPS type</th>
<th>Protein (mg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTM3R11</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>Ropy</td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>VTM3R12</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>Ropy</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>VTM3R17</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>NR</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>VTM4R20</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>VTM1R25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>Ropy</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>VTM3R22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>VTM3R24</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>VTM1R29</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>VTM3R32</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>~</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>VTM1R35</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>Ropy</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>VTM1R38</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>Ropy</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>VTM1R41</td>
<td>~</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>Ropy</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>VTM3R42</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>Ropy</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>VTM2R45</td>
<td>+</td>
<td>-</td>
<td>~</td>
<td>+</td>
<td>+</td>
<td>Ropy</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>VTM4R46</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>B/-</td>
<td>Ropy</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>VTM2R47</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NR</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>VTM1R48</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>Ropy</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>VTM1R50</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>Ropy</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>VTM2R55</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>Ropy</td>
<td>0.69</td>
<td></td>
</tr>
</tbody>
</table>

Streptococcus galloyticus subsp. macedonicus

<table>
<thead>
<tr>
<th>Isolates</th>
<th>Ornithine</th>
<th>Histidine</th>
<th>Tyrosine</th>
<th>Lysine</th>
<th>Lipase</th>
<th>Simmons (b/s)</th>
<th>EPS type</th>
<th>Protein (mg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTM1R33</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>NR</td>
<td>n/a</td>
</tr>
</tbody>
</table>

190
<table>
<thead>
<tr>
<th></th>
<th>+</th>
<th>+</th>
<th>-</th>
<th>+</th>
<th>+/-</th>
<th>Ropy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VTM2R39</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td></td>
<td>0.63</td>
</tr>
<tr>
<td>VTM4R51</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+B/+B</td>
<td>NR</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus alcis</em> sp. nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTM1R28</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+B/+</td>
<td>Ropy</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus vermontensis</em> sp. nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTM3R19</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>NR</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus gallolyticus</em> subsp. <em>mannosilyticus</em> subsp. nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTM3R15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>~/-</td>
<td>NR</td>
<td>n/a</td>
</tr>
<tr>
<td>VTM1R31</td>
<td>+</td>
<td>~</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.38</td>
</tr>
<tr>
<td>VTM1R44</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>Ropy</td>
<td>0.27</td>
</tr>
<tr>
<td>VTM4R49</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+B/+B</td>
<td>NR</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus gallolyticus</em> subsp. <em>melibiosilyticus</em> subsp. nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTM3R23</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>Ropy</td>
<td>0.72</td>
</tr>
<tr>
<td>VTM3R37</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+B/+B</td>
<td>Ropy</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Streptococcus gallolyticus</em> subsp. <em>ruminantium</em> subsp. nov.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VTM1R53</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td>NR</td>
<td>n/a</td>
</tr>
<tr>
<td>VTM4R54</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-/-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total isolates</td>
<td>(+)</td>
<td>31</td>
<td>15</td>
<td>28</td>
<td>34</td>
<td>25</td>
<td>15</td>
</tr>
</tbody>
</table>
Table 6-2 Acid production and ability of isolates to produce an acid byproduct from a variety of carbohydrates.

Acid production in skim milk was measured as pH over time. Positive acid production results (+) were indicated by a color change. All isolates produced acid from D-fructose, galactose, D+glucose, glycerol, lactose, maltose and sucrose (data not shown).

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Acid production (pH)</th>
<th>Acid production from carbohydrates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6hr</td>
<td>24hr</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N-Acetylglucosamine</td>
</tr>
<tr>
<td>VTM3R11</td>
<td>5.83</td>
<td>4.72</td>
</tr>
<tr>
<td>VTM3R12</td>
<td>6.14</td>
<td>5.11</td>
</tr>
<tr>
<td>VTM4R13</td>
<td>6.35</td>
<td>5.34</td>
</tr>
<tr>
<td>VTM1R14</td>
<td>6.13</td>
<td>5.11</td>
</tr>
<tr>
<td>VTM3R17</td>
<td>6.11</td>
<td>5.38</td>
</tr>
<tr>
<td>VTM4R20</td>
<td>5.95</td>
<td>4.70</td>
</tr>
<tr>
<td>VTM3R21</td>
<td>6.15</td>
<td>5.07</td>
</tr>
<tr>
<td>VTM3R22</td>
<td>5.88</td>
<td>4.69</td>
</tr>
<tr>
<td>VTM3R24</td>
<td>5.91</td>
<td>4.71</td>
</tr>
<tr>
<td>VTM1R25</td>
<td>6.04</td>
<td>4.84</td>
</tr>
<tr>
<td>VTM3R26</td>
<td>6.12</td>
<td>5.12</td>
</tr>
<tr>
<td>VTM1R27</td>
<td>5.57</td>
<td>4.25</td>
</tr>
<tr>
<td>VTM1R29</td>
<td>5.95</td>
<td>4.61</td>
</tr>
<tr>
<td>VTM3R32</td>
<td>6.01</td>
<td>4.96</td>
</tr>
<tr>
<td>VTM1R35</td>
<td>6.13</td>
<td>4.92</td>
</tr>
<tr>
<td>VTM1R38</td>
<td>6.25</td>
<td>4.80</td>
</tr>
<tr>
<td>VTM1R41</td>
<td>5.84</td>
<td>4.47</td>
</tr>
<tr>
<td>VTM3R42</td>
<td>5.79</td>
<td>4.84</td>
</tr>
<tr>
<td>VTM2R45</td>
<td>6.07</td>
<td>4.74</td>
</tr>
<tr>
<td>VTM4R46</td>
<td>6.05</td>
<td>5.57</td>
</tr>
<tr>
<td>VTM2R47</td>
<td>5.97</td>
<td>4.44</td>
</tr>
<tr>
<td>VTM1R48</td>
<td>5.91</td>
<td>4.61</td>
</tr>
<tr>
<td>Sample</td>
<td>Streptococcus gallolyticus subsp. macedonicus</td>
<td>Streptococcus alcis sp. nov.</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>VTM1R50</td>
<td>6.08</td>
<td>5.52</td>
</tr>
<tr>
<td>VTM2R55</td>
<td>5.88</td>
<td>4.82</td>
</tr>
<tr>
<td>VTM1R33</td>
<td>6.07</td>
<td>4.98</td>
</tr>
<tr>
<td>VTM2R39</td>
<td>5.88</td>
<td>4.57</td>
</tr>
<tr>
<td>VTM4R51</td>
<td>6.08</td>
<td>5.52</td>
</tr>
<tr>
<td>VTM1R28^t</td>
<td>5.88</td>
<td>4.64</td>
</tr>
<tr>
<td>VTM3R19^t</td>
<td>6.11</td>
<td>5.09</td>
</tr>
<tr>
<td>VTM3R15</td>
<td>6.18</td>
<td>5.85</td>
</tr>
<tr>
<td>VTM1R31^t</td>
<td>5.96</td>
<td>4.61</td>
</tr>
<tr>
<td>VTM1R44</td>
<td>6.08</td>
<td>5.30</td>
</tr>
<tr>
<td>VTM4R49</td>
<td>6.10</td>
<td>4.56</td>
</tr>
<tr>
<td>VTM3R23</td>
<td>5.98</td>
<td>5.08</td>
</tr>
<tr>
<td>VTM3R37^t</td>
<td>6.06</td>
<td>5.85</td>
</tr>
<tr>
<td>VTM1R53</td>
<td>6.22</td>
<td>6.21</td>
</tr>
<tr>
<td>VTM4R54^t</td>
<td>6.14</td>
<td>5.67</td>
</tr>
<tr>
<td>Total (+) isolates</td>
<td>29</td>
<td>33</td>
</tr>
</tbody>
</table>
Table 6-3 G+C content of isolates and DNA-DNA hybridization to reference strains Streptococcus gallolyticus gallolyticus (ATCC 700065) and S. gallolyticus macedonicus (ATCC BAA-249).

DNA-DNA hybridization (DDH) was calculated based on ΔTm between reference and hybrid strains, and regression equations generated from previously published DHH data for S. gallolyticus gallolyticus (R2=0.734) and S. gallolyticus macedonicus (R2=0.481) (Schlegel at el., 2003).

<table>
<thead>
<tr>
<th>Isolate</th>
<th>% G+C Content</th>
<th>S. gallolyticus gallolyticus</th>
<th>S. gallolyticus macedonicus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% G+C Content</td>
<td>ΔTm °C</td>
<td>% DDH</td>
</tr>
<tr>
<td>VTM1R28T</td>
<td>S. alcis sp. nov.</td>
<td>38.0</td>
<td>1.0</td>
</tr>
<tr>
<td>VTM3R19T</td>
<td>S. vermontensis sp. nov.</td>
<td>37.9</td>
<td>1.2</td>
</tr>
<tr>
<td>VTM1R31T</td>
<td>S. gallolyticus mannosilyticus subsp. nov.</td>
<td>37.9</td>
<td>1.9</td>
</tr>
<tr>
<td>VTM3R37T</td>
<td>S. gallolyticus melibiosilyticus subsp. nov.</td>
<td>38.1</td>
<td>2.0</td>
</tr>
<tr>
<td>VTM1R54T</td>
<td>S. gallolyticus ruminantium subsp. nov.</td>
<td>38.1</td>
<td>2.4</td>
</tr>
</tbody>
</table>
Supplemental Figure 1  Growth of isolates at varying temperatures (A), salinities (B) and pH (C), measured as absorbance at 600 nm.
Supplemental Figure 2 Protein production of ropy-strain isolates, measured as protein precipitate at 750 nm. Casein standards were used to generate a standard curve ($R^2=0.995$).

Supplemental Figure 3 Acid production of isolates in skim milk media incubated at 37°C. The pH of cultures was measured at 1, 6 and 24 hours to determine acid production over time. Initial pH of the media was calculated using a sterile blank as control.
CHAPTER 7    FIBROLYTIC BACTERIA ISOLATED FROM THE RUMEN OF NORTH AMERICAN MOOSE (ALCES ALCES).

Suzanne L. Ishaq*,1, Doug Reis2, and André-Denis G. Wright1,3

1 Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, Burlington, Vermont 05405

2 Department of Microbiology and Molecular Genetics, College of Agriculture and Life Sciences, University of Vermont, Burlington, Vermont 05405

3 Present address: School of Animal and Comparative Biomedical Sciences, College of Agriculture and Life Sciences, University of Arizona, Tucson, Arizona 85721

*Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. slpelleg@uvm.edu. Fax 1-802-656-8196.

Keywords: Bacillus/bacteria/fibrolytic/moose/probiotic
7.1 Abstract

Fibrolytic bacteria were isolated from the rumen of North American moose (*Alces alces*), which eat a high-fiber diet of woody browse. Thirty-one isolates were cultured from moose rumen digesta samples collected in Vermont. Using Sanger sequencing of the 16S rRNA gene, culturing techniques, and optical densities, isolates were identified and screened for biochemical properties important to plant carbohydrate degradation. The 31 isolates had the following percent identities to known sequences in the NCBI database: *Bacillus licheniformis*, 98–100% (n=22); *B. foraminis*, 98% (n=1); *B. firmus*, 98% (n=1); *B. flexus*, 100% (n=1); *B. niabensis*, 98% (n=1); *Paenibacillus woosongensis*, 98% (n=1); and *Staphylococcus saprophyticus*, 99–100% (n=4). Isolates were able to digest cellulose (n=31), cellobiose (n=28), xylan (n=26), starch (n=21), carboxymethylcellulose (n=21), and lignin (n=18) under minimal nutritional conditions. Fifteen isolates were able to digest all six carbohydrates or plant components tested. Isolates were able to tolerate up to 10% (n=16) salinity, between pH 4.0 (n=27) and pH 10.0 (n=27), and between 20°C (n=28) and 55°C (n=30). Isolates were tolerant to sodium azide (n=30), could reduce potassium tellurite (n=3), metabolize mannitol (n=29), produce indole from tryptophan (n=4), and all isolates could use citrate or propionate as a sole carbon source, as well as ammonium ions for nitrogen.
7.2 Introduction

Fibrolytic bacteria in the digestive tract of ruminants are instrumental in the digestion of plant matter for the host. The North American moose (*Alces alces*) is a large cervid, which consumes a high-fiber diet of woody browse: mainly willow, pine, maple, and fir (Belovsky and Jordan, 1981; Shipley, 2010). They also consume seasonally available aquatic vegetation, which is higher in sodium than arboreal vegetation (Belovsky and Jordan, 1981). This diet provides several nutritional challenges for which the moose has adapted. Moose produce tannin-binding salivary proteins to reduce the digestibility-reducing effects of tannins (Austin et al., 1989), and have a large liver: body size which may help them detoxify secondary metabolites found in willow and conifers (Shipley, 2010). Species of rumen bacteria that are resistant to secondary metabolites have been identified in some ruminants (Odenyo and Osuji, 1998; Dailey et al., 2008; Sundset et al., 2008), but have not yet been described in moose.

Few studies have identified the rumen bacteria of moose (Ishaq and Wright, 2012, 2014), or used culturing techniques to isolate bacteria from the rumen of moose (Dehority, 1986). Previously, it was shown that moose from Vermont contained a higher proportion of bacteria belonging to the phylum Firmicutes, which are mostly fibrolytic (Ishaq and Wright, 2014).

199
7.3 Methods

Fresh rumen samples were collected during the October 2010 hunting season in Vermont, with permission of licensed hunters through the Vermont Department of Fish and Wildlife. Hunters were given written and verbal instructions to collect samples from well inside the rumen, and to seal the container quickly to reduce oxygen exposure. Whole rumen samples (i.e. fluid and particulate matter) collected during field dressing were frozen within 2 h of death, and were transferred to the laboratory within 24 h, where they were mixed with an equal volume of 80% glycerol and stored at -80°C until culturing. Additional information regarding the hosts can be found in Ishaq & Wright (2012). Isolates were given unique identifiers (i.e. VTM3R11) containing the following abbreviations: Vermont (VT), moose (M), individual number (1–4), and rumen (R), as well as isolate number.

Bacteria were isolated on M8 agar plates (Bryant and Robinson, 1961; Dehority and Grubb, 1976), with an added 2 g/L of cellulose and cellobiose, inside an anaerobic chamber (COY Laboratories, Michigan, US). Whole rumen contents were serial diluted, and all dilutions (10^-1 to 10^-9) were plated with five replicates. Plates were monitored for up to 7 d, and colonies were picked and re-isolated on fresh media until colonies were shown to be pure using gram staining and colony morphology measurements. A total of 31 isolates were cultured from four individual moose rumen samples, and stock aliquots of each isolate were stored at -80°C. Isolates were tested for their catalase reaction (Gordon et al., 1973).
Monocultures were identified using automated cycle sequencing at the University of Vermont DNA Analysis Facility. The bacterial 16S rRNA gene was amplified using the universal bacterial primers 27F and 1494R (Lane, 1991). PCR was performed using the iTaq DNA Polymerase kit (Bio-Rad, California, US) following the manufacturer’s instructions. PCR conditions were: initial denaturation of 94°C for 5 min, then 33 cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min, followed by a final extension of 72°C for 6 min. PCR was performed on a C1000 thermal cycler (Bio-Rad, California, US). Recalcitrant isolates were first extracted using the DNA extraction protocol in the QIAamp DNA Stool Mini Kit (QIAGEN, Maryland, USA). Sequences were proofread using ChromasPro ver. 1.7.5, and aligned using the CLUSTALW algorithm in MEGA ver. 6.0. The alignment was visually inspected, and then used to calculate pairwise genetic distance using the Kimura 2-parameter model (Tamura et al., 2011). Sequences were classified using BLAST (NCBI) and compared to published sequences of fibrolytic bacteria from NCBI, and a neighbor joining tree was generated using MEGA.

As cellulose in the broth media prevented accurate optical density measurements, isolates were subcultured into 10 ml of tryptic soy broth (TSB) (1% vol/vol inoculation), and then incubated for 24 h at various temperatures or pH (adjusted prior to autoclaving). Optical density was used to determine relative growth using a Spectronic 200 (ThermoScientific, California, US), with absorbance measured 600 nm (Georgalaki et al., 2002). All samples were run in triplicate, and optimal ranges were set
as all isolates measuring >0.5% absorbance. Optimal salinity was measured as growth on tryptic soy agar (TSA) media with 4-15% NaCl. Heat tolerance was tested by immersing 48 h old cultures in a 60°C water bath for 30 min, then inoculating TSA plates (1% vol/vol inoculation) and incubating at 37°C for up to 72 h to observe for growth. Isolates which were able to survive >55°C were tested for their ability to tolerate sodium azide. Isolates were grown on azide dextrose media (tryptone, 15g/L; beef extract, 4.5g/L; glucose, 7.5g/L; sodium chloride, 7.5g/L; and sodium azide, 0.2g/L; pH 7.2) and incubated at 45°C for 5 d and observed for growth.

Isolates were tested for their ability to digest complex carbohydrates (cellulose, cellobiose, carboxymethylcellulose, xylan, and starch) or plant components (lignin) on minimal media (Bandounas et al., 2011). Minimal media plates were incubated at 37°C for up to two weeks to observe for growth. Isolates were tested on mannitol media for their ability to metabolize mannitol and tolerate potassium tellurite. To test for the production of the aromatic compound indole from the amino acid tryptophan, isolates were grown in 1% w/v tryptone broth for 14 d, after which Kovac’s reagent was added to the culture broth to test for a color reaction (Kovacs, 1928). Isolates were subcultured into Simmon’s Citrate slants (Simmons, 1926) and Propionate Slants (Gordon et al., 1973) for 7 d, and observed for bacterial growth and color change to indicate the ability to use citrate or propionate, respectively, as a carbon source and ammonia as a nitrogen source (Georgalaki et al., 2000). To test the ability to reduce nitrate to nitrite, isolates
were subcultured into nitrate broth for 2 and 7 d and then tested for color change using potassium iodine strips moistened with 1N HCl.

7.4 Results

All 31 isolates were gram positive and catalase positive. Isolates had the following percent identity to known sequences in NCBI: *Bacillus licheniformis*, 98–100% (n=22); *B. foraminis*, 98% (n=1); *B. firmus*, 98% (n=1); *B. flexus*, 100% (n=1); *B. niabensis*, 98% (n=1); *Paenibacillus woosongensis*, 98% (n=1); and *Staphylococcus saprophyticus*, 99–100% (n=4) (Table 1, Figure 1). All 16S rRNA sequences are available from NCBI (KP245773- KP245803).

All 31 isolates tolerated 4% NaCl (data not shown). Isolates (n=16) were able to tolerate up to 10% salinity (Table 1), including *B. firmus*, *B. flexus*, some *B. licheniformis*, *B. niabensis*, *P. woosongensis*, and some *S. saprophyticus*. Isolates from all species grew to >0.5 absorbance between pH 4.0 (n=27) and pH 10.0 (n=27), and between 20°C (n=28) and 55°C (n=30) (Figure 2). Twenty-nine isolates exhibited excellent growth after heat shock, but two isolates (*B. niabensis* VTM4R58, and *S. saprophyticus* VTM2R99) exhibited no growth. All but one *B. licheniformis* isolate (VTM3R64) tolerated sodium azide and exhibited growth after 5 d.

Under minimal conditions, isolates were able to digest cellobiose (n=28), xylan (n=26), starch (n=21), carboxymethylcellulose (n=21), and lignin (n=18) (Table 1). All 31 isolates were able to grow on cellulose, glucose, and lactose (data not shown), and 15
isolates were able to digest all six carbohydrates tested (Table 1). Twenty-seven isolates were able to metabolize mannitol and produce a color change, but four *B. licheniformis* could not (VTM2R66, VTM1R71, VTM1R80, VTM1R88). Only two *B. licheniformis* isolates (VTM2R66, VTM2R82) and one *B. foraminis* isolate (VTM4R85) could reduce tellurite. Two *B. licheniformis* isolates (VTM1R74, VTM1R75), one *B. firmus* isolate (VTM2R84), and one *B. foraminis* isolate (VTM4R85) were able to produce indole from tryptophan. All isolates were able to use citrate and propionate as their carbon source, and use ammonia for nitrogen. Twelve isolates were able to reduce nitrite to nitrate after 48 h, and an additional two isolates (*S. saprophyticus* VTM2R99 and *B. firmus* VTM2R84) were able to reduce nitrite to nitrate after 7 d of growth (Table 1).

### 7.5 Discussion

Thirty-one fibrolytic bacterial isolates were examined for their biochemical capabilities and potential as a probiotic for ruminants. Based on their ability to survive a wide range of growth parameters and digest complex carbohydrates even on minimal media, many of the thirty-one fibrolytic isolates in the present study have the potential for use in agricultural or industrial applications. However, the ability to survive in the developing digestive tract using milk or milk replacer as a substrate, as well as the ability to consistently grow well in culture, are also important considerations for a viable probiotic product.
Bacillus licheniformis is an important member of the rumen community as it produces a variety of extracellular enzymes which can digest lignocelluloses (Archana and Satyanarayana, 1997), starches (Saito, 1973; Pen et al., 1992), keratin (Lin et al., 1992), and acetate (Veith et al., 2004). It is able to digest glucose anaerobically (Veith et al., 2004), and certain strains show antibiotic resistance (Pollock, 1965; Moews et al., 1990). The industrial applications of B. licheniformis are extensive due to the breadth of its enzymatic products, but also because many are thermophilic or halophilic (Gordon et al., 1973; Veith et al., 2004; Wei et al., 2010). The present study identified 22 isolates which had greater than 98% sequence identity to B. licheniformis, as well as four isolates, which had greater than 98% sequence identity to B. foraminis, B. firmus, B. flexus, and B. niabensis.

Staphylococcus saprophyticus is a cellulolytic bacterium originally isolated from the termite gut (Paul et al., 1986). The present study identified four isolates which had greater than 99% sequence identity to S. saprophyticus. Paenibacillus woosongensis was originally isolated from forest soil, and was shown to digest a variety of carbohydrates, including cellulose and xylan (Lee and Yoon, 2008). One isolate in the present study had a 98% sequence identity to P. woosongensis.

The present study found that 15 out of 31 isolates were able to digest all six different carbohydrates and plant components investigated on minimal media. Lignin is an aromatic alcohol polymer found in the cell walls of plants and some algae. In the cell wall it is often bonded to cellulose or hemicelluloses, which increases the durability of
plants cell walls, but decreases their digestibility. After cellulose, lignin is the second most abundant polymer on Earth. Xylans are hemicelluloses which are found in the cell wall of plants, especially hardwoods, and some algae, and if it is not fermented by gut microorganisms it can decrease the absorption of minerals in the intestines (Jiang K, 1986). Mannitol is a sugar found in species of deciduous flowering ash. Carboxymethylcellulose is a purified form of cellulose that is more soluble in water, thus it is used in food production, pharmaceuticals, or industrial applications.

The ability to survive under restrictive nutritional conditions is especially important trait for bacteria used in industrial applications, but can also provide an advantage over competitive species or strains of bacteria which require vitamins or other substrates in the rumen. Additionally, bacteria which can positively impact the host, would be beneficial to overall animal health in addition to increasing dietary efficiency. Indole production often takes place in the intestines, and is used as a quorum-sensing signal molecule between gut bacteria. However, its presence in the intestines also stimulates cellular junction-associated molecules in gut epithelial cells, and promotes resistance to dextran sodium sulfate (DSS)-induced colitis (Shimada et al., 2013).

7.6 References


206


7.7 Figures

Figure 7-1 Phylogenetic comparison of 31 isolates which known sequences (NCBI).
A neighbor-joining tree was created using MEGA ver. 6 and the Kimura 2-parameter model.
Figure 7-2 Growth at various temperatures (A) and pHs (B), as measured by optical density at 600 nm.
### 7.8 Tables

Table 7-1 Isolate GenBank ID, closest GenBank match with percent identity, growth on minimal media or on high salinity, and ability to reduce nitrite.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>GenBank ID</th>
<th>Closest GenBank Identification</th>
<th>CMC</th>
<th>Cellobiose</th>
<th>Lignin</th>
<th>Starch</th>
<th>Xylan</th>
<th>8% NaCl</th>
<th>10% NaCl</th>
<th>Nitrate reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTM4R85</td>
<td>KP245773</td>
<td>98% <em>B. foraminis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R84</td>
<td>KP245774</td>
<td>98% <em>B. firmus</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R86</td>
<td>KP245775</td>
<td>100% <em>B. flexus</em></td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM4R61</td>
<td>KP245776</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R62</td>
<td>KP245777</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM4R63</td>
<td>KP245778</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM3R64</td>
<td>KP245779</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R65</td>
<td>KP245780</td>
<td>98% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R66</td>
<td>KP245781</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM2R67</td>
<td>KP245782</td>
<td>100% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM4R68</td>
<td>KP245783</td>
<td>98% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM4R69</td>
<td>KP245784</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM4R70</td>
<td>KP245785</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM1R71</td>
<td>KP245786</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM1R72</td>
<td>KP245787</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R73</td>
<td>KP245788</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R74</td>
<td>KP245789</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R75</td>
<td>KP245790</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM3R76</td>
<td>KP245791</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM3R77</td>
<td>KP245792</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM3R78</td>
<td>KP245793</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM1R80</td>
<td>KP245794</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R81</td>
<td>KP245795</td>
<td>99% <em>B. licheniformis</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R82</td>
<td>KP245796</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM1R88</td>
<td>KP245797</td>
<td>99% <em>B. licheniformis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM4R58</td>
<td>KP245798</td>
<td>98% <em>B. niabensis</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM1R92</td>
<td>KP245799</td>
<td>98% <em>P. woosongensis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>VTM1R96</td>
<td>KP245800</td>
<td>100% <em>S. saprophyticus bovis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM4R98</td>
<td>KP245802</td>
<td>100% <em>S. saprophyticus bovis</em></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>VTM4R97</td>
<td>KP245801</td>
<td>99% <em>S. saprophyticus s</em></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VTM2R99</td>
<td>KP245803</td>
<td>99% <em>S. saprophyticus s</em></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Total positive (n=31) | 21 28 18 21 26 19 16 14

B = Bacillus; P = Paenibacillus; S = Staphylococcus
CHAPTER 8  THE EFFECTS OF ADMINISTERING A FIBROLYTIC PROBIOTIC MADE FROM MOOSE RUMEN BACTERIA TO NEONATAL LAMBS.

Suzanne L. Ishaq\textsuperscript{1,*}, Christina J. Kim\textsuperscript{1}, and André-Denis G. Wright\textsuperscript{1,2}

\textsuperscript{1} Department of Animal Science, College of Agriculture and Life Sciences, University of Vermont, 570 Main St., Burlington, Vermont 05405

\textsuperscript{2} Present address: School of Animal and Comparative Biomedical Sciences, College of Agriculture and Life Sciences, University of Arizona, Tucson, Arizona 85721

* Suzanne Ishaq, Department of Animal Science, University of Vermont, 203 Terrill Building, 570 Main Street, Burlington VT 05405. spelleg@uvm.edu Fax) 1-802-656-8196.

Keywords/Cross-reference: 16S rRNA gene, lambs, moose, probiotic
8.1 Abstract

The present study investigated the effect of a fibrolytic probiotic, using bacteria isolated from the rumen of the North American moose (*Alces alces*), and which was administered daily to neonate lambs until 1 week after weaning. It was hypothesized that regular administration of a fibrolytic probiotic to neonate animals through weaning would increase the developing rumen bacterial diversity, increase animal production, and allow for long-term colonization of the probiotic species. Neither weight gain nor wool quality was improved in lambs given a probiotic, but dietary efficiency was increased as evidenced by the reduced feed intake (and rearing costs) without a loss to weight gain. Additionally, the probiotic lambs had a lower acetate to propionate ratio than control lambs, which has previously been shown to indicate increased dietary efficiency. Sampling coverage was high in the first two time points, after which it decreased. Conversely, Shannon, Inverse Simpson, CHAO, and ACE were low and increased over time, all of which is a function of the increasing diversity of the rumen microbiota as the rumen develops. The experimental group had a higher diversity at the beginning of the experiment. Fibrolytic bacteria made up the majority of sequences. In all time points and both groups, *Prevotella* was the most prevalent genus, while *Butyrivibrio* and *Ruminococcus* were also prevalent. While protozoal densities increased over time and were stable, methanogen densities varied greatly in the first six months of life for lambs. This is likely due to the changing diet and bacterial populations in the rumen.
8.2 Introduction

Over the first few months of life, the rumen microbiome of the neonate ruminant undergoes rapid shifts as the animal weans and changes diets, and the rumen develops. As the rumen develops, it increases in size until it is the largest stomach chamber, its rumen papillae become longer and more differentiated, and a stable microbiota is established. Initially, lactic acid bacteria (LAB), such as *Streptococcus thermophilus*, *Lactobacillus acidophilus*, or *Bifidobacterium bifidus*, tend to dominate, as well as *Escherichia coli* (Fonty et al., 1987; Minato et al., 1992). While cellulolytic bacteria do appear in the rumen within the first few days of life (Fonty et al., 1987; Minato et al., 1992; Morvan et al., 1994), it is not until weaning and a transition to a plant-based diet that they become the dominant type of rumen bacteria (Sinha and Ranganathan, 1983; Ishaq and Wright, 2014). As the microbial diversity adapts to the rumen environment and the diet provided, so, too, do the gastrointestinal tract epithelia adapt to the microbiota. Host cells can eventually recognize conserved cell-surface microbial markers, and the presence of these will active pro- or anti-inflammatory responses, as well as host epithelial cell signaling (Chang, 2008; Abreu, 2010). Thus, introducing new microbiota after these host-microbiota interactions have been made may not be successful.

Different weaning practices can influence the developing microbiota. Including a creep feed or hay along with a milk diet can encourage larger populations of fibrolytic bacteria
(Yáñez-Ruiz et al., 2010), improve starch and fiber digestion (Poe et al., 1971), increase volatile fatty acid production (especially acetate and butyrate) (Laarman et al., 2012), and improve rumen development (Norouzian et al., 2011). Likewise, development of the rumen microbiota can be encouraged through early colonization by administering a probiotic. Probiotics for livestock are generally comprised of LAB or fibrolytic bacteria. LAB probiotics are more common in pre-weaned ruminants (Abe et al., 1995; Vlková et al., 2010; Ripamonti et al., 2011) or monogastrics (Abe et al., 1995; Pajarillo et al., 2015), but are also used in adult ruminants (Aikman et al., 2011; Boyd et al., 2011). Fibrolytic probiotics have more often been used to improve digestive function in adult ruminants (Kumar and Sirohi, 2013; Præsteng et al., 2013), as well as for pre-weaned ruminants (Sun et al., 2010). Many studies report short-term beneficial effects only, either due to the production animals reaching market weight, or because the probiotic failed to colonize the digestive tract long-term.

It was hypothesized that regular administration of a fibrolytic probiotic to neonate animals through weaning would increase the developing rumen bacterial diversity, increase animal production, and allow for long-term colonization of the probiotic species. The present study investigated the effect of a fibrolytic probiotic, which had been created using bacteria isolated from the rumen of the North American moose (*Alces alces*), and which was administered daily to neonate lambs until 1 week after weaning at 9 weeks of age. Neonatal ruminants undergo rumen development over a period of 8-12 weeks,
during which the rumen and reticulum increase in size and functionality (Hobson and Fonty, 1997), making the weaning period an ideal time period for rumen manipulation.

Moose were chosen as the source for probiotic strains as they are highly likely to host efficient species or strains of bacteria, which can digest cellulose, hemicellulose, and lignin. Moose subsist on a diet of woody browse which is very high in fiber (Belovsky, 1981; Molvar et al., 1993; Routledge and Roese, 2004). Additionally, their body temperature and dry matter intake is more similar to lambs than to calves or goat kids, thus improving the likelihood of long-term rumen colonization by the species of interest (Gasaway and Coady, 1974; Franzmann et al., 1984; Piccione et al., 2003; Dwyer and Morgan, 2006; Committee on the Nutrient Requirements of Small Ruminants, 2007; Kochan, 2007; Piccione et al., 2007).

8.3 Methods

All procedures were approved by the UVM Institutional Animal Care and Use Committee (protocol (14)-008). Results are presented by date and experimental week (week).

Twenty Dorset-cross lambs, 4-7 days of age, were purchased from Bonnieview Farm, Craftsbury, VT. Lambs were group housed at the Miller Research Farm at the University of Vermont (UVM), Burlington VT, beginning on April 22, 2014. Eighteen male and
two female lambs were randomly assigned to either the Control (n=10) or the Experimental (n=10) group with nine males and one female per group. Males were castrated within the first two weeks of the study, and groups had similar weights (mean 5.9±0.2 kg) prior to the beginning of the study. Water was provided ad libitum. For four weeks, lambs were fed DuMOR lamb milk replacer (Tractor Supply Co, Shelburne VT) using bucket feeding systems (Premier 1 Supplies, Washington, IO), and group intake was recorded. Beginning in week five, lambs were also given DuMOR sheep starter pelleted grain feed (Tractor Supply Co, Shelburne VT), and group intakes were recoded. At week six, lambs were weaned off of the milk replacer and were fed grain pellets and timothy hay, and again group intake was recorded. At eight weeks (June 26, 2014) lambs were transferred to Sterling College in Craftsbury, VT, where they were maintained as a single mob grazing on pasture until mid-October, 2014.

8.3.1 Probiotic

Five bacterial isolates were chosen for use as a probiotic based on a previous study (Ishaq, Reis, et al., 2015). Isolates are as follows, with GenBank accessions numbers in parentheses: Bacillus foraminis VTM4R85 (KP245773), B. firmus VTM2R84 (KP245774), B. licheniformis VTM2R66 (KP245781), B. licheniformis VTM1R74 (KP245789), and Staphylococcus saprophyticus bovis VTM1R96 (KP245800). Isolates were selected based on their ability to digest carboxymethylcellulose, cellobiose,
cellulose, lignin, starch, and xylan on minimal media. Isolates were also able to survive at a wide range of temperatures, salinities, and pH.

Isolates were cultured separately in M8+cellulose broth for approximately six months to determine whether the isolate could be maintained for an extended period at sufficient concentrations to be used as a probiotic. Purity was determined via weekly gram staining, and occasional Sanger sequencing as previously described (Ishaq, Reis, et al., 2015). Concentration was measured by number of colony forming units (CFUs) on a plate count, performed in duplicate. As per Food and Drug Administration (FDA) regulations, probiotics must maintain $10^7$ CFUs for the duration of its shelf life. The five isolates were then tested for their ability to survive in commercial milk replacer for up to 72 hr at 37°C, and maintain a minimum density of $10^7$. Isolates were cultured for 24, 48 and 72 hr in DuMOR Blue Ribbon lamb milk replacer (DuMOR 06-9551-0234), reconstituted according to manufacturer instructions, and then replated on M8+cellulose for plate counts at 24 hr.

Isolates were grown individually in M8 broth supplemented with cellulose as previously described (Ishaq, Reis, et al., 2015). Cultures were checked regularly for purity using gram staining, and concentrations were measured using standard plate counts. Twenty-four hour old cultures were combined at equal concentration within 1 hour of administration and kept cool during transport. One ml of inoculant or blank media was
administered orally to experimental and control lambs, respectively, daily between noon and 1 pm. After two weeks, when lambs were approximately 20 days old, the dose was increased to 2 ml/day. Probiotic or blank media was given daily for 9 weeks until weaning at 9.5 to 10 weeks of age.

8.3.2 Production

Lambs were weighed every 2-3 days for the first three weeks, then weekly through weaning, then monthly for the duration of the study. At study week 8, when lambs were weaned and put on pasture, a 2 x 2 in patch was shaved on the side, within 3-5 in of the spine (i.e. mid-side sample). Wool was allowed to grow out for 14 weeks, after which a 1 x 1 in patch was shaved, dried, weighed, and sent for fiber testing to Yocom-McColl Testing Labs in Denver, CO. Significance for this and other measurements was calculated using Student’s T-test, and deviation is presented as standard deviation (SD) or standard error mean (SEM).

8.3.3 Rumen sampling

Rumen samples were collected weekly for eight weeks, then monthly once lambs were on pasture. Samples were collected in the morning, prior to weaning this was within one to two hours of feeding. Esophageal tubing was used to obtain samples directly from the rumen, from which up to 15 ml of fluid and particulate matter (ruminal contents) were collected and put on ice immediately. Some rumen fluid was separated out and used to
measure pH and volatile fatty acids. Rumen pH was tested using a MW101 pH meter (Milwaukee, NC). Volatile fatty acids were measured using gas chromatography at the William H. Miner Agricultural Research Institute (Chazy, NY). Thawed rumen samples were centrifuged for 20 min at 4° at 10,000 x G. Supernatant was filtered through a single layer of Whatman filter paper, and 0.8 ml of filtrate was mixed with an equal volume of internal standards (oxalic acid and trimethylamine).

8.3.4 Sequencing and DNA data analysis

DNA was extracted from individual samples using the QIAamp DNA stool fast kit (QIAGEN, MD), and the V1-V3 region of the 16S rRNA gene was amplified using previously described protocols (Ishaq and Wright, 2014). Amplicons were sent to Molecular Research, LP (MR DNA) in Shallowater, TX for MiSeq ver. 4.

Sequences were analyzed using MOTHUR ver. 1.31 (Schloss et al., 2009; Kozich et al., 2013). Sequences were trimmed to remove barcodes and primers, as well as any sequence that contained a mismatch in the barcode, more than two mismatches in the primer, sequences with homopolymers >8, sequences < 475 bases or >570 bases, and sequences with an average quality score <32 over 5 bases. Sequences were aligned to the Silva 16S rRNA bacteria MOTHUR reference file, which had been modified to include fibrolytic isolates cultured in the laboratory, including the five which were used in the probiotic. The reference alignment was also trimmed to begin at 27F and end after 800
bases. Chimeras were identified using UCHIME (Edgar et al., 2011) and removed. Sequences were identified using the k nearest neighbor method. Data were subsampled to 10,000 sequences per sample, clustered using the nearest neighbor method, and diversity parameters were measured. ACE (Chao and Shen, 2003), CHAO (Chao and Shen, 2010), Good’s Coverage (Good, 1953), Shannon-Weiner diversity (Shannon and Weaver, 1949), AMOVA and Unifrac values are presented as group mean.

In order to compare control and experimental groups from all four time points, sequences which passed QA were pooled, and were subsampled to 2,000 sequences per sample, giving 20,000 per group per time point. This subsample was used to create a neighbor-joining tree using the mother-integrated algorithms for Clearcut (Evans et al., 2006), linear discriminant analysis (LDA) using the mother-integrated Lefse (Segata et al., 2011), and principal component analysis (PCoA).

8.3.5 Real-time PCR

Real-time PCR was used to calculate archaeal and protozoal densities in whole samples. DNA was amplified using a CFX96 Real-Time System (Bio-Rad, CA) and a C1000 ThermalCycler (Bio-Rad, CA). Data were analyzed using CFX Manager Software ver. 1.6 (Bio-Rad, CA). The iQ SYBR Green Supermix kit (Bio-Rad, CA) was used: 12.5µL of mix, 2.5µl of each primer (40mM), 6.5µL of ddH₂O, and 1µL of the initial DNA extract diluted to approximately 10 ng/µL. For methanogens, the primers targeted the
methyl coenzyme-M reductase A gene (mcrA), following the protocol by Denman et al. (2007). The internal standards for methanogens were a mix of *Methanobrevibacter smithii, M. gottschalkii, M. ruminantium* and *M. millerae* (R²=0.998). For protozoa, the primers, PSSU316F and PSSU539R (Sylvester et al., 2004), targeted the 18S rRNA gene, following the protocol by Sylvester et al. (2004), and The internal standards for protozoa were created in the laboratory using fresh rumen contents as previously described (Ishaq et al., unpublished). Both protocols were followed by a melt curve, with a temperature increase 0.5°C every 10s from 65°C up to 95°C to check for contamination.

### 8.4 Results

#### 8.4.1 Probiotic

Five isolates (VTM2R66, VTM1R74, VTM2R84, VTM4R85, VTM1R96), which were selected for further testing, maintained concentrations ranging from 10⁷ to 10¹⁰ CFUs over six months. The same five isolates were able to maintain densities greater than 10⁷ in liquid lamb replacer over 72 hours.

#### 8.4.2 Production

Total group weight was higher in the control group for nearly the duration of the study, with the exception of the week 8 weighing; however, this time point was the only one which came close to statistical significance (p=0.06). The total cost at the end of weaning (week 8) of milk replacer, starter grain, and timothy hay for the experimental
groups was $1564.63, at a weaning weight of 248.95 kg for the group. This gives a yield of 6.28 $/kg. The total cost of the control group was $1592.17, and at a weaning weight of 263.25 kg this yields 6.05 $/kg. When taking into account the total group weights at market weight (aged six months, week 23), the cost/group drops to 6.08 $/kg for the experimental group and increases to 6.19 $/kg for the control group.

Mid-side sample wool weight was higher (p=0.02) in the experimental group (mean=0.83 g, SEM=0.5) than in the control group (mean=0.67 g, SEM=0.07). Mean fiber diameter (MFD) was not significantly different (p=0.14) between experimental (MFD=34µ, SEM=0.6, SD=7.4) and control (MFD=33.1µ, SEM=0.6, SD=7.7) groups. The experimental group did have a significantly (p=0.04) lower coefficient of variation (CoV=21.8) than the control group (CoV=23.4). Although the experimental group had a higher percentage of fibers that were >30µm (66.8%, SEM=3.4 experimental, 62.1%, SEM=2.7 control), this was not statistically significant (p=0.11).

The average pH over the course of the experiment was 7.2 for the experimental group and 7.0 for the control group (Figure 2). The experimental group had a higher average pH for the first seven weeks of the experiment and lower variability within the group, while the control group were more likely to have a higher average for the remainder of the study. Seven out of 12 sampling time points were significantly different (Figure 2).
Total volatile fatty acids were not significantly different between Experimental and Control lambs (Figure 3); however, groups were significantly different at weeks 5, 11 and 23 when comparing total VFAs including ethanol. Total VFAs were highest at weeks 8 and 15, and lowest at week 9 after being on a hay only diet for one week (Figure 3). Acetate, propionate, and butyrate were significantly (p<0.05) higher in experimental lambs at weeks 15 and 23, while lambs were on pasture. The acetic acid to propionic acid ratio was statistically lower in the experimental group at weeks 9, 11 and 15 (Figure 4).

8.4.3 Sequencing

Between 11,000 and 95,000 unique sequences passed quality assurance steps per sample, giving a total of 500,000 to 1.1 million sequences per data set of 20 samples. At the first sampling time point, week 2, after administering the probiotic for a week, there was little statistical difference in rumen diversity between groups, except for AMOVA, and unweighted and weighted Unifrac (Table 1). At week 6, CHAO, ACE, Shannon, and Coverage were different between groups, with higher diversity in the experimental groups. Groups were not statistically different on any diversity measure except for unweighted Unifrac at week 11 in July, after being on pasture for two weeks (Table 1). However, by week 23 in October, the control group showed higher diversity according to Shannon and Inverse Simpson, and groups clustered separately by weighted and unweighted Unifrac.
When comparing all time points together in the smaller subsampled data set, there were 1,787 OTUs identified from the 160,000 sequences, with 88 (4.9%) discriminatory OTUs with LDA >2 (p<0.05). The number of discriminatory OTUs were as follows: week 2, five experimental, four control; week 6, five experimental, six control; week 11, 37 experimental, 17 control; and week 23, nine experimental, five control. PCoA graphs showed a strong clustering of groups by sampling time (Figure 5), but not by treatment (data not shown), indicating that the change of rumen bacteria over the course of rumen development is a stronger indicator of variance.

Bacteroidetes was the most prevalent phylum (38-73% of total sequences) in both groups for the duration of the study, with the exception of the first sampling of the control group (Figure 6). Firmicutes was the second most prevalent (23-59%). In the control group, Bacteroidetes increased while Firmicutes decreased for the first three sampling (weeks 2, 6 and 11), while the experimental had a general trend of decreasing Bacteroidetes and increasing Firmicutes over time. Other prominent phyla tended to peak at one or two time points, including Actinobacteria and Proteobacteria (week 6), Fibrobacteres (week 11), and Synergistetes (weeks 11, 23).

Major families identified are shown in Figure 6, with families belonging to Bacteroidetes as shades of blue and members of Firmicutes in shades of green. Prevotellaceae (mostly
species *Prevotella*) was a prominent family in all time points and in both groups, but was significantly higher in the experimental group at week 2. Lachnospiraceae was also prominent in all samples, although it was significantly higher in the control group at week 2. The experimental group had more Ruminococcaceae than the control group at weeks 2 and 6. There were also families which were prominent in only one time point, such as Bacteroidaceae, Streptococcaceae, and the candidate family p-2534-18B5 in week 2; Coriobacteriaceae (mostly species *Olsenella*) in week 6, Fibrobacteraceae in week 11, and the candidate Family XI of the class Bacilli (phylum Firmicutes) in week 23.

While the genera *Bacillus* and *Staphylococcus* were found in both groups at all time points, there was not enough resolution in the sequencing amplicons to accurately identify probiotic sequences down to species or strain. The total genera identified were as follows: week 2, 301 experimental, 273 control; week 6, 183 experimental, 184 control; week 11, 292 experimental, 331 control; and week 23, 482 experimental, 483 control (Supplemental Material). Overall, 694 genera were identified across both groups and all time points. The most prevalent genus in all groups and time points was *Prevotella*. Other prominent genera included *Bacteroides, Butyrivibrio, Catabacter, Clostridium, Dialister, Lactobacillus, Olsenella, Orbacterium, Parvimonas, RC9, Ruminococcus, Selemonas, and Streptococcus* (Supplemental Material).
8.4.4  Real-time PCR

Protozoal density in both control and experimental groups increased over time until they leveled off at approximately $2 \times 10^3$ (Figure 8). Control group had statistically higher (p<0.05) densities at weeks 8 and 23. Methanogen densities increased for the first month, then rapidly decreased at week 6 (Figure 8). Levels peaked at week 8, at which point densities decreased to week 11, then peaked again at week 15. While average density was higher in control lambs for most time points, due to the variability of densities within groups this was not statistically significant. In lambs with elevated protozoal densities, methanogen density was also elevated (Figure 9). When protozoal densities were low, methanogen densities were also low. However, this trend was not statistically significant ($R^2=0.376$)

8.5  Discussion

Neither weight gain nor wool quality was improved in lambs given a probiotic, however, dietary efficiency was increased as evidenced by the reduced feed intake (and rearing costs) without a significant loss to weight gain. This reduction in rearing costs would be further amplified using more traditional husbandry practices, such as rearing lambs outside and giving them access to grass during weaning, thus precluding the need to supplement with hay. Additionally, the probiotic lambs had a lower acetate to propionate ratio than control lambs, which has previously been shown to indicate increased dietary
efficiency (Van Soest, 1982; Morgavi et al., 2012). An increased production of propionate reduces free hydrogen in the rumen, making it less available to methanogens.

Sampling coverage was high in the first two time points, after which it decreased. Conversely, Shannon, Inverse Simpson, CHAO, and ACE were low and increased over time, all of which is a function of the increasing diversity of the rumen microbiota as the rumen develops. While there were differences between groups in terms of statistical diversity, there was no difference in OTUs/sample between groups, and this is likely due to evenly subsampling the data set. The experimental group had a higher diversity at the beginning of the experiment, but this was not persistent.

Fibrolytic bacteria made up the majority of sequences identified in samples, and while some fibrolytic genera were elevated in the experimental group, this was not consistent across all time points. In all time points and both groups, *Prevotella* was the most prevalent genus, while *Butyrivibrio* and *Ruminococcus* were also prevalent. All three genera have previously been shown to be fibrolytic (Smith et al., 1973; Maglione et al., 1992; Daniel et al., 1995; Cotta and Forster, 2006; Suen et al., 2011). However, sequences from the probiotic strains could not be confidently reported in the sequenced samples.
While protozoal densities increased over time and were stable, methanogen densities varied greatly in the first six months of life for lambs. This is likely due to the changing diet and bacterial populations in the rumen. For example, when methanogen density decreased at week 6, the proportion of acetogenic bacteria increased (i.e. phylum Actinobacteria, and species such as *Acetivibrio* and *Acetitomaculum* in the phylum Firmicutes, class Clostridia), fostering competitive pathways to methanogenesis (Lopez et al., 1999). Reducing methanogenesis would not only make for more eco-friendly livestock, but would also reduce the amount of energy lost to the host which might have otherwise been used for production.

Despite the small increase in dietary efficiency, a more dramatic increase in production might result from altering the probiotic administered in the present study. Increasing the dosage, using a different mix of fibrolytic bacteria, or using a probiotic with fibrolytic and lactic-acid bacterial strains, are all potential methods of improving upon the results presented here.

### 8.6 Acknowledgements

The authors would like to thank Matt Bodette, Laura Cersosimo, Sam Frawley, Emma Hurley, Anjana Mangalat, Katy Nelligan, Scott Shumway, Sam Rosenbaum, Lee Warren, and Sarah Zegler for their assistance with animal husbandry and sample collection, Louise Calderwood for facilitating the transfer of lambs from UVM to Sterling College,
Mike Richards for care of sheep at Sterling College, and Dr. Sabrina Greenwood for assistance with UVM Miller Farm facilities used for this project.

8.7 References


Ishaq, S. L., M. A. Sundset, J. Crouse, and A.-D. G. Wright. 2015. High-throughput DNA sequencing of the moose rumen from different geographical location reveals a core ruminal methanogenic archaeal diversity and a differential ciliate protozoal diversity. unpublished.


8.8 Figures

Figure 8-1 Group weight (kg) means over time. Significance is denoted with *, and error bars show standard error mean.
Figure 8-2 Group pH means over time.
Significance is denoted with *, and error bars show standard error mean.
Figure 8-3 Volatile fatty acid (VFA) and ethanol profile of groups (E=experimental, C=control) for two time points on pasture.
Figure 8-4 Acetate to propionate ratio, with SEM.

Figure 8-5 Principal component analysis of samples by sampling time: May=teal, June=green, July=red, and October=dark blue.
Figure 8-6 Diversity at the phylum level for all sequencing time points.
Error bars show standard error mean.
Figure 8-7 Diversity at the family level for all sequencing time points.
Error bars show standard error mean.
Figure 8.8 Real-time PCR data for methanogens and protozoa.
Significance is denoted with *, and error bars show standard error mean.
Figure 8-9 Comparison of methanogen versus protozoal densities for all samples
Table 8-1 Diversity statistics per sample for each of the four sampling time points. Results are listed by group, Experimental (n=10) and Control (n=10), or All (n=20).

<table>
<thead>
<tr>
<th></th>
<th>Group</th>
<th>5-1-14</th>
<th>6-4-14</th>
<th>7-10-14</th>
<th>10-1-14</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total sequences which passed QA steps</strong></td>
<td>All</td>
<td>953,581</td>
<td>1,002,520</td>
<td>501,909</td>
<td>1,007,093</td>
</tr>
<tr>
<td><strong>Subsampled to 10,000/sample (200,000/time point)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good’s Coverage</td>
<td>Exp</td>
<td>0.85</td>
<td>0.83 *</td>
<td>0.63</td>
<td>0.32 *</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>0.87</td>
<td>0.87</td>
<td>0.62</td>
<td>0.28</td>
</tr>
<tr>
<td>Shannon-Weiner Diversity</td>
<td>Exp</td>
<td>3.23</td>
<td>3.69 *</td>
<td>5.82</td>
<td>7.86 *</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>3.25</td>
<td>3.09</td>
<td>5.72</td>
<td>8.10</td>
</tr>
<tr>
<td>Inverse Simpson</td>
<td>Exp</td>
<td>6.34</td>
<td>8.02</td>
<td>21</td>
<td>220 *</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>6.35</td>
<td>6.15</td>
<td>19</td>
<td>447</td>
</tr>
<tr>
<td>CHAO</td>
<td>Exp</td>
<td>10,907</td>
<td>11,489 *</td>
<td>50,674</td>
<td>95,992</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>8,513</td>
<td>7,712</td>
<td>57,141</td>
<td>111,950</td>
</tr>
<tr>
<td>ACE</td>
<td>Exp</td>
<td>31,864</td>
<td>31,738 *</td>
<td>149,728</td>
<td>278,155</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>22,578</td>
<td>21,378</td>
<td>171,878</td>
<td>352,438</td>
</tr>
<tr>
<td><strong>Total species-level OTUs</strong></td>
<td>All</td>
<td>20,706</td>
<td>21,035</td>
<td>70,062</td>
<td>106,823</td>
</tr>
<tr>
<td><strong>Mean species-level OTUs/sample</strong></td>
<td>Exp</td>
<td>1,271</td>
<td>1,425</td>
<td>3,918</td>
<td>5,700</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>1,156</td>
<td>999</td>
<td>3,977</td>
<td>6,190</td>
</tr>
<tr>
<td><strong>Shared OTUS (sequences)</strong></td>
<td>All</td>
<td>9 (65,731)</td>
<td>7 (68,258)</td>
<td>20 (64,589)</td>
<td>19 (19,215)</td>
</tr>
<tr>
<td><strong>Group shared OTUs (sequences)</strong></td>
<td>Exp</td>
<td>13 (72,661)</td>
<td>12 (65,582)</td>
<td>38 (96,060)</td>
<td>31 (78,967)</td>
</tr>
<tr>
<td></td>
<td>Con</td>
<td>12 (77,050)</td>
<td>12 (63,501)</td>
<td>37 (95,910)</td>
<td>33 (81,424)</td>
</tr>
<tr>
<td><strong>AMOVA (p-value)</strong></td>
<td>All</td>
<td>0.01 *</td>
<td>0.73</td>
<td>0.30</td>
<td>0.14</td>
</tr>
<tr>
<td><strong>Weighted Unifrac</strong></td>
<td>All</td>
<td>0.80 *</td>
<td>0.47</td>
<td>0.44</td>
<td>0.57 *</td>
</tr>
<tr>
<td><strong>Unweighted Unifrac</strong></td>
<td>All</td>
<td>0.87 *</td>
<td>0.67 *</td>
<td>0.77 *</td>
<td>0.87 *</td>
</tr>
</tbody>
</table>

*Denotes statistically significant value (p<0.05) between groups at that time point.
### 8.10 Supplemental Material

This table has been modified from the original, which is too large to include here. It was modified to include the top genera by abundance.

<table>
<thead>
<tr>
<th>taxon</th>
<th>Experimental week 2 %</th>
<th>Control week 2 %</th>
<th>Experimental week 6 %</th>
<th>Control week 6 %</th>
<th>Experimental week 11 %</th>
<th>Control week 11 %</th>
<th>Experimental week 23 %</th>
<th>Control week 23 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acholeplasma</td>
<td>0.0</td>
<td>0.0</td>
<td>0.6</td>
<td>0.3</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>Acinetobacter</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Actinobacillus</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Aggregatibacter</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Alysiella</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Anaeroplasma</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>1.2</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Anaerostipes</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Anaerovorax</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Bacillus</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>3.7</td>
<td>13.6</td>
<td>0.1</td>
<td>0.1</td>
<td>1.2</td>
<td>1.0</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Bergeriella</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Blautia</td>
<td>0.1</td>
<td>2.3</td>
<td>1.1</td>
<td>1.9</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Butyribrio</td>
<td>0.2</td>
<td>0.1</td>
<td>3.1</td>
<td>0.5</td>
<td>2.8</td>
<td>1.5</td>
<td>1.2</td>
<td>2.6</td>
</tr>
<tr>
<td>Butyribrio-Pseudobutyribrio</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Catabacter</td>
<td>0.1</td>
<td>2.9</td>
<td>0.2</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Clostridium</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>Coprococcus</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Desulfovibrio</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Dialister</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.7</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>Dorea</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Erysipelothrix</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Eubacterium</td>
<td>0.5</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>1.0</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Fastidiosipila</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>Species</td>
<td>Value1</td>
<td>Value2</td>
<td>Value3</td>
<td>Value4</td>
<td>Value5</td>
<td>Value6</td>
<td>Value7</td>
<td>Value8</td>
</tr>
<tr>
<td>-------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Filifactor</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Fusobacterium</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Geospirorobacter</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Guggenheimella</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Helcococcus</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Hydrogenoanaerobacterium</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Incertae_Sedis</td>
<td>2.2</td>
<td>4.3</td>
<td>6.4</td>
<td>3.6</td>
<td>1.6</td>
<td>1.7</td>
<td>3.2</td>
<td>4.1</td>
</tr>
<tr>
<td>Johnsonella</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>Lactobacillus</td>
<td>1.8</td>
<td>4.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Mannheimia</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Mogibactera</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Mogibactera</td>
<td>0.1</td>
<td>0.4</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Moraxella</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Morvella</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Neisseria</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Olsenella</td>
<td>0.0</td>
<td>0.0</td>
<td>11.2</td>
<td>10.4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Oribacterium</td>
<td>0.0</td>
<td>0.0</td>
<td>9.7</td>
<td>10.5</td>
<td>0.2</td>
<td>0.4</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Oscillibacter</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Oscillospira</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.6</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Parabacteroides</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Parvimonas</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>6.4</td>
<td>2.4</td>
</tr>
<tr>
<td>Pedobacter</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Peptostreptococcus</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Porphyromonas</td>
<td>0.8</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Prevotella</td>
<td>33.3</td>
<td>13.7</td>
<td>35.4</td>
<td>38.2</td>
<td>42.4</td>
<td>42.9</td>
<td>31.9</td>
<td>32.9</td>
</tr>
<tr>
<td>RC9</td>
<td>0.7</td>
<td>0.5</td>
<td>1.8</td>
<td>0.6</td>
<td>5.5</td>
<td>4.4</td>
<td>6.9</td>
<td>6.6</td>
</tr>
<tr>
<td>Roseburia</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Ruminococcus</td>
<td>0.1</td>
<td>0.0</td>
<td>2.4</td>
<td>2.6</td>
<td>1.8</td>
<td>1.1</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Selenomonas</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.6</td>
<td>0.4</td>
<td>8.8</td>
<td>5.7</td>
</tr>
<tr>
<td>Soehngenia</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Solobacterium</td>
<td>0.2</td>
<td>0.9</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>2.8</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Spirochaeta</td>
<td>1.2</td>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Streptococcus</td>
<td>0.2</td>
<td>9.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Subdoligranum</td>
<td>0.3</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Succinicolasticum</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Tannerella</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Treponema</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>0.6</td>
<td>0.1</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>5.9</td>
<td>2.6</td>
<td>0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>vadinBC27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victivallis</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.1</td>
<td>0.2</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>Xylanibacter</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
</tr>
</tbody>
</table>
CHAPTER 9 CONCLUSION

9.1 Summary

It was hypothesized that the moose would host novel microorganisms; that these microorganisms would be capable of a wide variety of enzymatic functions; and that these microorganisms could be used to improve other systems. The objectives of this work were to identify the bacteria (CHAPTERS 2, 3), archaea (CHAPTER 5), and protozoa (CHAPTERS 4, 5) present in the rumen of moose; to culture bacteria from the rumen of the moose in order to study their biochemical capabilities (CHAPTERS 6, 7); and to apply those cultured bacterial isolates to improve fiber digestion in neonate lambs (CHAPTER 8).

9.2 The moose rumen: summarized findings and unanswered questions

Despite the extensive work presented here on the moose rumen microbiota, it cannot be said that the moose rumen is yet fully understood. A large proportion of bacteria and protozoa taxa studied here could not be identified, indicating that there are yet more novel taxa which need to be isolated, cultured, classified, and deposited into reference databases. This is particularly necessary for the protozoa, since relatively few species have been cultured and identified. As much of that work was done prior to the emergence of DNA sequencing technology, and given the difficulty in maintaining protozoa in culture, even fewer of the known GIT species have 18S rRNA sequences.
available in public databases (currently <250 sequences). In addition, fungi have barely been investigated in the moose GIT [1], and the bacteriophages not at all.

Aside from the literature and work presented here on SSU rRNA, no work has been done on the moose rumen microbiome: the enzymes at work, the myriad beneficial or toxic products being produced, and the wealth of undocumented microbial genetic material present. The bacteria in the rumen of the moose are predominantly fibrolytic, as they should be given the diet of moose. However, is this fibrolytic dominance seasonally tied to diet quality, forage content, and cellulose intake, as speculated and shown previously [2–4], or is the fibrolysis in fact accomplished by different bacterial species in different locations regardless of season. The genus *Prevotella* has been shown to be amylolytic and fibrolytic, but without increased sensitivity of identification it remains to be seen which species or subspecies is present, and which digestive function those found in the Alaskan moose were performing. Targeted sequencing of known genes which code for fibrolytic enzymes, or shotgun sequencing of the entire rumen contents to target full-genomes, would reveal the enzymatic potential of the rumen as a whole. Known and putative enzymes could be identified from DNA sequencing or microarrays, and the functional microbiome can be studied using RNA sequencing or microarrays. Previously, in vitro [5] or in vivo [6, 7] work was used to study digestibility of plant matter by the moose rumen microorganisms.
In the work presented here, methanogens in the moose rumen were identified for the first time; although actual methane production in moose was not studied, either in vitro or in vivo. It was speculated here that moose might have a lower production of methane than other ruminants due to the high proportion of forage in their diet [8] and faster rate of food passage through their GIT [9]. Previously, an in vitro study of biogas production from GIT methanogens showed that moose rumen microorganisms produced less biogas than those from beaver [10]. An ability to efficiently digest cellulose leads to a reduction in methanogenesis [11]; however, the digestion of starch, also can reduce methanogenesis via two means. Firstly, the production of propionate is increased which sequesters free H and prevents its use in methanogenesis [12]. Second, the increase in amylolytic bacteria reduces the pH, which in turn reduces the protozoal population, again providing less H for methanogenesis [12]. Thus, it is likely that moose produce less methane than other ruminants, although it is possible that this may be location/diet specific.

Methane production from moose can be most easily studied using anaerobic culturing, either of pure methanogen cultures or whole rumen contents. One method would be to use a tabletop methane digester, but any containment system which allowed for the introduction of plant biomass and the collection of biogas could be used. An in vivo study could be performed using methane collection systems such as a respiration chamber, combined feeder/gas measurement systems, or by estimating methane output by
measuring the output of a tracer gas [13]. Given the size of moose and the relative few numbers of captive moose available, this would seem impractical.

9.3 Fibrolytic probiotics in lambs: beneficial or bust?
The administered probiotic did not produce the expected results; however, it cannot be considered a failed hypothesis. Weight gain was not improved in experimental lambs over control lambs, with the exception of a period after starter grain was introduced in which experimental lambs gained 30% body weight over the previous week, which was the highest percent gain at any point in either group. This may be an indication that the rumen and rumen microbiota of experimental lambs were more developed, and were able to take advantage of a solid food diet more quickly than those of control lambs. Rumen papillae and intestinal villi length measurements were not able to taken; however, they would have provided insight into whether the GIT of probiotics lambs were, in fact, more developed. This might have been performed using GIT epithelial samples for traditional histological slides or scanning with micro-computed tomography (micro-CT) [14] to measure length and differentiation. When put in the context of feeding costs per kg of body weight; however, the lower cost of raising the probiotic lambs without sacrificing weight gain is indicative of an increased dietary efficiency.

Although total wool growth was significantly higher in experimental lambs, wool quality was not improved over that of control lambs. Wool quality is determined by staple
(fiber) length, strength, diameter, and diameter uniformity. Thinner diameter wool (<22 microns) can be woven into a thinner, finer yarn, and thus is more valuable. The sheep used in this study were varieties of Dorset-crosses. Dorset is a dual-purpose breed, producing good quality meat and medium quality (avg. 27-33 microns) white wool. On the farm from which the lambs were obtained, sheep were mainly bred for meat and dairy (cheese) production, although wool was also collected and sold. It may be that any positive effects of the probiotic were not enough to overcome the genotypic determination of wool quality.

The probiotic lambs had a higher and more stable pH than control lambs through weaning (Figure 1). A stable pH is important when transitioning between diets, especially when transitioning to a highly digestible starch, such as grain, to prevent a low rumen pH which can kill rumen microorganisms and disrupt digestive function. Sheep are not prone to rumen acidosis, unless they gain access to a high-starch feedstuff and gorge themselves. However, dairy cattle are more sensitive to rumen acidosis, which can decrease production and may require medical intervention (SECTION 1.3.5). As the lambs in the probiotic study were acting as a model for cattle, a stable rumen pH would be an important benefit of the probiotic.

It is difficult to determine whether the probiotic strains properly colonized the rumen or not, as the amplicons used for analysis did not have the resolution to be identified past
genus or, in some cases, family. Difficulty identifying Illumina sequences down to genus level has previously been seen [15]. In the present work, a variety of reference databases (i.e. GenBank, RDP, Silva) were used, as well as multiple classification methods (i.e. k nearest neighbor, wang [16]) and confidence cutoff levels, in an effort to classify down to species and subspecies, but to no avail. The genera *Bacillus* and *Staphylococcus* were identified, but in low numbers in both groups. The probiotic strains may have colonized the rumen of experimental lambs, validating the hypothesis that the lamb rumen diversity could be altered long-term, but without providing the anticipated beneficial effects.

It is possible that the dosage of bacteria was not high enough to allow for larger-scale colonization of the rumen. It may also be the case that the isolates tested were not ideal for colonization in some way which was not investigated beforehand (0). For example, isolates may be sensitive to antimicrobial products released by other rumen microorganisms, or may have been subject to predation by protozoa. To improve the identification of probiotic strains, sequencing using longer amplicons or with another platform could be attempted to improve the resolution. Another method would be to create strain-specific real-time PCR primers and look at 16S copy number in whole samples. PCR product would have to be amplified using additional cycles in order to generate a detectable signal if there are very low copy numbers. Potentially, whole rumen samples could be used to reisolate the probiotic strains in culture; however, this method would be the least efficient.
Probiotic strains were shown to survive in milk replacer for a period of time, which allows for the potential of administering during bulk-feeding. The probiotic was not tested as a top-dressing; either as a live culture or dried cells, but this represents another possible means of administering the probiotic to many animals simultaneously. However, despite the benefits provided by the probiotic as outlined previously, the probiotic is not a viable product in its current form. To be useful to producers, a significant economic difference in animal product quality or production cost must be shown. Though a reduction in production (rearing) cost was shown, it is not enough to compensate for the estimated cost of adding a probiotic product. A more dramatic increase in production might result from altering the probiotic administered in the present study. Increasing the dosage, using a different mix of fibrolytic bacteria, or using a probiotic with fibrolytic and lactic-acid bacterial strains are all potential methods of improving upon the results presented in CHAPTER 8. Additionally, probiotic benefits may be more pronounced when administered to dairy cattle, or when also measuring milk production. Immunological measurements, such as inflammation in the gut, GIT cytokines, blood cortisol, blood cholesterol, or meat quality measurements, such as concentrations of fatty acids, could also be used to detect beneficial changes caused by the probiotic other than a direct increase in production.
9.4 Molecular methodology

9.4.1 Sequencing Platform

High-throughput sequencing was chosen as the platform for the main work presented in this dissertation, as the previous investigations into the moose rumen had been using anaerobic culturing or light microscopy [17–21]. Initially, preliminary investigations into the rumen and colon bacteria of the moose were made using DNA microarray chips [22]. It quickly became clear that this did not provide the sensitivity to detect specific genera of interest, nor was it capable of identifying unknown taxa. The DNA chip was able to quickly determine broad differences in diversity between samples, and in the years since the DNA microarray experiment was performed in 2011 [22], the chip has been redesigned to detect 50,000 bacterial taxa using over a million different probes.

Multiplexed sequencing-by-synthesis was the next obvious choice, using barcoded primers to sequence multiple samples simultaneously; however, there were multiple platforms to choose from. In 2011, when the first data sets were being considered for sequencing, Roche 454 was the more appropriate platform. It was capable of producing de novo sequences up to 400-500 bases in length, with tens of thousands of sequences generated per sample. To perform pyrosequencing, the Roche platform uses a cyclic flow of nucleotides over a plate containing wells, with one strand of ssDNA per well. When a nucleotide triphosphate matched the candidate sequence it would hybridize via
DNA polymerase. This would release pyrophosphate, which is converted by ATP sulfurylase to ATP and used by luciferase to convert luciferin to oxyluciferin.

This reaction produces a measurable amount of light, which is measured by a camera and converted to a “flow value”, and is outputted as a flowgram. Flow values do not directly translate to a specific number of nucleotides, thus it was necessary to use bioinformatics algorithms such as PyroNoise [23] to determine the number of bases in the flow using maximum likelihood. This provided an extra step of quality assurance, as candidate flowgrams can be compared to reference flowgrams generated from mock data sets and adjusted to remove noise [23, 24].

Aside from Roche, the most promising platform being developed was Illumina, which at the time boasted 25-100 bases, with millions of sequences generated. The Illumina platforms were faster, as four different fluorescent dyes were used to tag nucleotide bases, thus negating the need to pause between read steps. Once the nucleotide base was incorporated, the dye was cleaved and diffused out of the read zone. However, the sequence length being generated was simply not long enough for taxonomic resolution. For large contig-assembly and full-genome sequencing, the shotgun sequencing approach of Illumina works well and continues to expand its potential [25].
The Illumina platforms rapidly evolved, and just a few short years later were able to produce amplicons of 400-500 bases using paired-end reads, although by this time Roche was still outpacing them at 800-1,000 base non-paired reads [26]. It also became more economical than Roche sequencing, as fewer reagents were required and sequencing preparation time was greatly reduced [27]. In late 2013, Roche announced that it was closing the 454 Life Sciences subsidiary and discontinuing the 454 reagents by 2016, further driving an industry-wide switch from Roche 454 to Illumina platforms.

The Illumina platforms came with their own drawbacks. Raw error rates were higher [27, 28], especially after the first 60 nucleotide bases [15], although consensus finished-read accuracy was 99% [29]. Data output came in the form of fasta files, as opposed to Roche which provided flowgrams that could undergo noise reduction. Additionally, Illumina sequencers showed lower coverage in high GC regions [30]. The sheer number of sequence reads produced created problems during analysis as random access memory (RAM) and hard drive space became limiting.

With respect to the data generated in this project, Roche 454 was the optimal platform for investigating bacterial 16S rRNA sequences and protozoal 18S rRNA sequences, and Illumina’s MiSeq platform was optimal for investigating archaeal 16S rRNA sequences. For example, when sequencing protozoa from three pooled Alaskan moose rumen samples using Roche 454, a total of 12,326 sequences passed QA steps, which were
represented by 769 unique (non-redundant) sequences [31]. When sequencing protozoa from these same three samples individually using MiSeq ver. 3, a total of 200,513 sequences passed the same QA steps, represented by 99,895 unique sequences [32]. This was a gross overestimation of protozoan diversity, and required more stringent analysis parameters to correct the data, such as using a higher quality score cutoff, or condensing sequences which had one or two polymorphic differences between them.

Likewise, when analyzing bacterial sequences generated from the lamb probiotic study, MiSeq ver. 3 generated over 1 million raw sequences per data set of 20 samples (CHAPTER 8). After stringent QA steps, between 500,000 and just over 1,000,000 total sequences were retained per data set, over 75% of which were considered unique. Again, this was an overestimation of diversity, but without the ability to reduce sequencing noise in the original sequences there was no way to detect raw sequencing errors. Additionally, computer RAM became limiting, and to accommodate such a massive amount of data needed to calculate genetic distance and cluster sequences into OTUs, it was necessary to subsample the data sets.

9.4.2 Sequencing Target

Properly selecting a sequencing target is extremely important as it impacts all aspects of the project: from PCR amplification, to sequencing reactions, to DNA sequence analysis. As previously mentioned (SECTION 1.2), the SSU rRNA gene was chosen as the target
for phylogenetic studies. As current high-throughput sequencing techniques were limited in the length of high-quality, low-error sequences which could be generated, a section of the rRNA gene was needed as a proxy for the full-length sequence. In addition to being the optimal size for high-throughput sequencing (~500 b), the amplicon needed to be easy to generate in the lab, have a trustworthy and expansive (when possible) reference database of sequences available for comparison [33–35], and not be liable to amplification bias by primers [15, 36, 37]. Most importantly, the amplicon needed to provide a similar statistical estimate of diversity and resolution for identification as when using the full-length gene [15, 38–40]. Thus, it was determined that for the presented work, the V1-V3 region of the 16S for bacteria and archaea, and the V3-V4 region of the 18S for protozoa, provided the overall best solution to the current constraints and the selected sequencing platform.

Most studies which investigate the microbiota of a certain environment tend to focus on one domain at a time, or employ different methods to investigate different taxa (i.e. bacteria, archaea, protozoa, and fungi). However, some phylogenetic studies attempt to use the same primers and amplification parameters across multiple kingdoms. While this may be a more efficient approach in terms of time, money, and effort in the laboratory, it is not an adequate means of investigation. For example, studies which use “universal prokaryotic primers” to co-amplify bacteria and archaea result in very few archaeal sequences as compared to the number of bacterial sequences generated [41–43]. In the
GI tract, archaeal density is only 1-2 logs lower than bacterial density, thus “universal” PCR parameters for the 16S gene overwhelmingly underrepresents archaea. Likewise, “universal eukaryotic primers” often marginalize protozoal sequences [42, 44]. In some instances, primers which were prokaryote-specific have been used to amplify eukaryotic 18S sequences [45, 46].

9.4.3 Bioinformatics Programs and Analysis

MOTHUR [24] was selected as the bioinformatics platform of choice as it integrated a comprehensive list of analytical functions within the same platform, was more user-friendly than other programs, and was compatible with Windows, Mac, and Linux operating systems. Notably, it integrated a version of PyroNoise [23] which reduced noise in pyrosequencing flowgrams generated from Roche 454 platforms, as well as a wide variety of chimera checking algorithms.

Two of the more popular chimera checking programs, UCHIME [47] and ChimeraSlayer [48], were compared using bacterial 16S rRNA sequences from the moose rumen generated using the Roche 454 platform (CHAPTER 3). UCHIME identified 18,146 chimeras out of a total 40,514 sequences, or 45% of sequences as chimeric, when using abundance to estimate chimeras. Using the same data set, again using abundance to estimate chimeras, ChimeraSlayer identified 17,225 chimeric sequences, or 43%. However, 14,933 out of the 18,146 UCHIME putative chimeric sequences (82%) were
able to classify with >80% confidence. For ChimeraSlayer, 14,268 out of the 17,225 putative chimeric sequences (83%) could be classified with >80% confidence. Using the Silva bacterial 16S reference file to estimate chimeric sequences reduced the percentage of sequences being identified as chimeric, as well as reduced the percentage of those which could actually be classified with a reasonable amount of confidence. Thus, UCHIME with a reference database was used for all further analyses.

Reference alignments for bacteria, archaea, and eukaryota from the Silva [49], Greengenes [50], and RDP [34] databases were provided for use with MOTHUR. The Silva database for bacteria was found to generate better alignments and classification than the RDP reference database (Table 1). Owing to the unique nature of the archaeal and protozoan sequences being investigated, and the low number of available rumen ciliate protozoan sequences in any database, it was necessary to generate in-house reference alignments for both using publically available sequences [31, 32].

9.5 References


42. Kittelmann S, Seedorf H, Walters WA, Clemente JC, Knight R: **Simultaneous amplicon sequencing to explore co-occurrence patterns of bacterial, archaeal and eukaryotic microorganisms in rumen microbial communities.** *PLoS One* 2013, **8**:e47879.


47. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R: **UCHIME improves sensitivity and speed of chimera detection.** *Bioinformatics* 2011, **27**:2194–2200.


9.6 Figures

Figure 9-1 Mean rumen pH over time, with standard deviation bars.
9.7 Tables

Table 9-1 A comparison of the Silva bacterial database vs the Ribosomal Database Project (RDP) bacterial reference files in ability to classify 16S rRNA gene bacterial sequences.

<table>
<thead>
<tr>
<th></th>
<th>Silva</th>
<th>Silva %</th>
<th>RDP</th>
<th>RDP %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Unique sequences</td>
<td>33,622</td>
<td></td>
<td>33,622</td>
<td></td>
</tr>
<tr>
<td>Total classified as Bacteria</td>
<td>33,622</td>
<td>100%</td>
<td>33,610</td>
<td>99.96%</td>
</tr>
<tr>
<td>Total classified as Archaea</td>
<td>0</td>
<td>0%</td>
<td>1</td>
<td>~0%</td>
</tr>
<tr>
<td>Unclassified at the Domain level</td>
<td>0</td>
<td>0%</td>
<td>11</td>
<td>0.03%</td>
</tr>
<tr>
<td>Total classified at the Phylum level (excl. Unclassified)</td>
<td>28,111</td>
<td>83.61%</td>
<td>26,769</td>
<td>79.6%</td>
</tr>
<tr>
<td>Total classified at the Family level (excl. Unclassified)</td>
<td>21,122</td>
<td>62.82%</td>
<td>16,277</td>
<td>48.41%</td>
</tr>
<tr>
<td>Total classified at the Genus level (excl. Unclassified)</td>
<td>10,642</td>
<td>31.66%</td>
<td>3,700</td>
<td>11.00%</td>
</tr>
<tr>
<td>Total classified at the Species level</td>
<td>0</td>
<td>0%</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Total Unclassified bacteria</td>
<td>5,511</td>
<td>16.39%</td>
<td>6,841</td>
<td>20.34%</td>
</tr>
</tbody>
</table>


283


Krause, D. O., R. J. Bunch, L. L. Conlan, P. M. Kennedy, W. J. Smith, R. I. Mackie, and C. S. McSweeney. 2001. Repeated ruminal dosing of Ruminococcus spp. does not result in persistence, but changes in other microbial populations occur that can be measured with quantitative 16S-rRNA-based probes. Microbiol 147:1719–1729.


MOTHUR. Available from: www.mothur.org


comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. BMC Genomics 13:1.


