APPLYING CUTTING EDGE DNA SEQUENCING TECHNOLOGY TO FURTHER OUR UNDERSTANDING ABOUT BOVINE HEALTH

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Svetlana Ferreira Lima

August 2017

APPLYING CUTTING EDGE DNA SEQUENCING TECHNOLOGY TO FURTHER OUR UNDERSTANDING ABOUT BOVINE HEALTH

Svetlana Ferreira Lima, Ph. D. (D.V.M.)

Cornell University 2017

Technical improvements in high-throughput sequencing technologies have opened new frontiers in microbiome research by allowing cost-effective characterization of complex microbial communities, including that of the bovine host. Targeted next generation sequencing has been shown to be an efficient approach for detection and identification of microorganisms and is more likely to be implemented in clinical and diagnostic settings due to its lower cost and shorter labor time. Such an approach relies on sequencing of a genetic marker, the 16S rRNA gene, for specific characterization of bacterial communities and bacterial pathogenic agents.

Given the potential role of the microbiome in animal health and disease, the overall objectives of this dissertation were to: 1) identify the most appropriate DNA extraction protocol that efficiently isolates a majority of the heterogeneous bacterial species encountered in non-mastitic and mastitic milk samples for accurate taxonomic profiling and detection of clinical mastitis causative agents (Chapter two); 2) use high-throughput sequencing of the 16S rRNA gene to characterize the bovine microbiome of distinct anatomical sites (mammary gland and upper respiratory tract) and its associations with bovine health (Chapters three and four); and 3) investigate the origin of the bovine microbiome (Chapter five). The present dissertation describes the effective application of the16S rRNA gene amplicon sequencing in clinical science and discovery of new pathogens. The knowledge gained from this research is novel and serves as an anchor for exploring the unfolding frontier of knowledge about the bovine microbiome across various anatomical sites and its effect on animal health.

BIOGRAPHICAL SKETCH

Svetlana (Svet) Ferreira Lima was born and raised in Goiânia, Goiás - Brazil. She received her degree in Veterinary Medicine in 2012 from the Veterinary Medicine and Animal Science School at the Federal University of Goiás, Brazil. During Veterinary School, under the mentorship of her ruminant theriogenology professor, Dr. Maria Lucia Gambarini, Svet undertook research projects that evaluated the molecular epidemiology, risk factors and immune response in dairy cows with vaginal infections due to *Ureaplasma diversum*. During this program Svet developed a strong appreciation for research. In 2011, by Dr. Gabarini's mediation, Svet was introduced to Dr. Rodrigo Bicalho who would later become her Ph.D. supervisor. In 2012, six months prior her graduation, Svet was invited to join Dr. Rodrigo Bicalho's research group at Cornell University as a visiting scholar. This experience augmented her appreciation for research with particular interest in dairy cattle health.

After the completion of her DVM program, she worked as a Relief Research Support Staff at the Preventive Medicine Department of Veterinary Medicine and Animal Science School (Federal University of Goiás /Brazil). During that period she participated in the expansion and consolidation of the Goiás State Support Center for diagnoses of poultry diseases. Although this program was not focused on dairy cattle health, it was essential for the improvement of her professional and research skills and an encouragement for pursuing a Doctor of Philosophy degree.

In 2013, Svetlana Lima was awarded with a scholarship from Science Without Borders, a program funded by the Brazilian government, and in the fall of 2013 she started her Ph.D. program in the field of Animal Science at Cornell University, minoring in microbiology and epidemiology, under the supervision of Dr. Rodrigo Bicalho.

ACKNOWLEDGMENTS

I would not have been able to complete this journey of both good and difficult times, lots of laughs and screams, accomplishments and frustrations, without the support of my family, friends and mentors. This dissertation is definitely a reflection of all the encouragement and guidance received during my doctorate.

To my loving family members, who have always motivated me to work hard and keep things in perspective. Thank you to my Mom (Jacqueline Lima) and Dad (Roberto dos Santos) for their love and sacrifices, you both have been and will always be my foundation. Many thanks to my siblings, Andressa Lima and Tales Mazzoccante, for all the love and also for the encouragement that you guys gave me to move abroad and pursue a dream. Nonetheless, I would like to thanks my grandparents, Luzia Lima and Otacilio Lima, for being the best grandparents a person could ask for.

My friends, from Cornell and elsewhere, who have showed me absolute support. Thanks to my friends Vinicius Machado, Rafael Neves, Eduardo Barbosa, Steven Moirano, Mariana Rosa, Marjory Xavier, Taryn Pena and Marilia Gomes for being there whenever I needed a friend. A special thanks to Bicalho's family Rodrigo, Marcela, Felipe and Rafael for being my "American/Brazilian family". Thanks to Marcela Bicalho for the constant support and motivation as well as for the occasional "suck it up and get it done" that was essential for the completion of my research program. I am also very much thankful to my previous Veterinary program supervisor, Maria Lucia, who has always been an example, good friend and guided me to Cornell University.

To my advisor Rodrigo Bicalho, thank you for being such an enthusiastic and great mentor. I have learned so much throughout my time here at Cornell mainly because Rodrigo has introduced me to so many different subjects and opportunities. Without his trust in me and his

patience I could not have completed this Ph.D. program. Also I would like to express my deepest sense of gratitude for my committee members Dr. McArt, Dr. Angert, and Dr. Huson.

Lastly, I would like to thanks the former and current members of the Bicalho' research group, in special Andre Gustavo Teixeira and the extern students, who have intellectually or physically supported my research. Many thanks to Hazel Higgins for always putting me in a good mood and for the support. Finally, I would like to express my sincere gratitude to the Science Without Border Program for funding my Ph.D. program.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADG, Average daily gain

ADL, Average days of life at first diagnosis

ANOSIM, Analysis of similarities

ANOVA, Analysis of variance

BC, Bovine colostrum

BRD, Bovine respiratory disease

BW, Birth weight

CM, Clinical mastitis

CMC, Clinical mastitic colostrum

C-section, Cesarean section

DNA, Deoxyribonucleic acid

DCC, Days carried calf

FDR, False discovery

IMI, Intra-mammary infection

LF, left front quarter

LH, left hind quarter

MRA, Mean relative abundance

NCMC, Non-clinical mastitic colostrum

OTU, Operational Taxonomic Unit

PCoA, Principal coordinate analysis

PCR, Polymerase chain reaction

QIIME, Quantitative Insights Into Microbial Ecology

RF, right front quarter

RH, right hind quarter

SAS, Statistical Analytical Software

SE, Standard error

SEM, Standard error of the mean

SD, Standard deviation QIIME, Quantitative Insights Into Microbial Ecology

STAMP, Statistical Analytical of Taxonomic and Functional Profile

URT, Upper respiratory tract

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

INTRODUCTION

Microbiome research

Since the discovery and elucidation of deoxyribonucleic acid (DNA) and its structure in the 1950s by Watson and colleagues (Watson and Crick et al., 1953, Feughelman et al., 1955, Wilkins et al., 1956) strategies and multiple technological innovations have been made in efforts to better understand the complexity and diversity of genomes in a variety of ecosystems, including those in the health and disease. The term metagenome was first described by Jo Handelsman in 1998 as "the genomes of the total microbiota found in nature" (Handelsman et al., 1998). More recently, with the development and application of high throughput sequencing methods in combination with the growing recognition of the importance of the microbial community a National Science and Technology Council Committee of US government scientists was created in 2015 (Stulberg et al., 2016). The US government scientists defined microbiome as "a multi-species community of microorganisms in a specific environment (that is, host, habitat or ecosystem)" and microbiome research as "those studies that emphasize community-level analyses with data derived from genome-enabled technologies" (Stulberg et al., 2016).

DNA SEQUENCING

Technology advances and current approaches

DNA-sequencing based technologies are vastly contributing to the expanding knowledge on microbiomes in the current ability to sequence millions of bases as well as the ability to sequence and assemble whole genomes (Metzker, 2010, Goodwin et al., 2016, Heather and Chain, 2016). The primordial Sanger sequencing technology was launched in 1977 by Frederick Sanger (Sanger et al., 1977) and was widely used for approximately 25 years until its replacement by second and third-generation sequencing methods (Di Bella et al., 2013). Sanger DNA sequencing technology, also commonly described as the chain termination method, is based on "selectively incorporation of chain-terminating dideoxynulceotides by the DNA polymerase during the in vitro DNA replication" (Sanger et al., 1977, Zhou X. and Li Y., 2015). In the laboratory procedures for clone library construction, the isolated DNA is "fragmented" and cloned into a vector for DNA amplification in bacterial cells, resulting in millions of individual bacterial colonies produced from one unique sample. The DNA from the clones is then individually isolated, followed by the sequencing reaction (Hoy, 2013). Although, the Sanger technology opened new avenues in genome research, it is still restricted in terms of the number of samples that are able to be processed and analyzed as well as the high cost per base of sequence.

The massively parallel sequencing complemented by the sequencing-by-synthesis design represents a state-of-the-art approach of the second-generation sequencing (e.g. 454 pyrosequencing, Illumina sequencing technologies: Hiseq and Miseq). This approach has been commercially available since 2005 (Di Bella et al., 2013) and was favored due to its reduced cost and high throughput (Loman et al., 2012). While differing in engineering structures as well as in sequencing chemistry, the next-generation sequencing platforms share the massive parallel sequencing approach and include barcoding and multiplexing of the samples (Di Bella et al.,

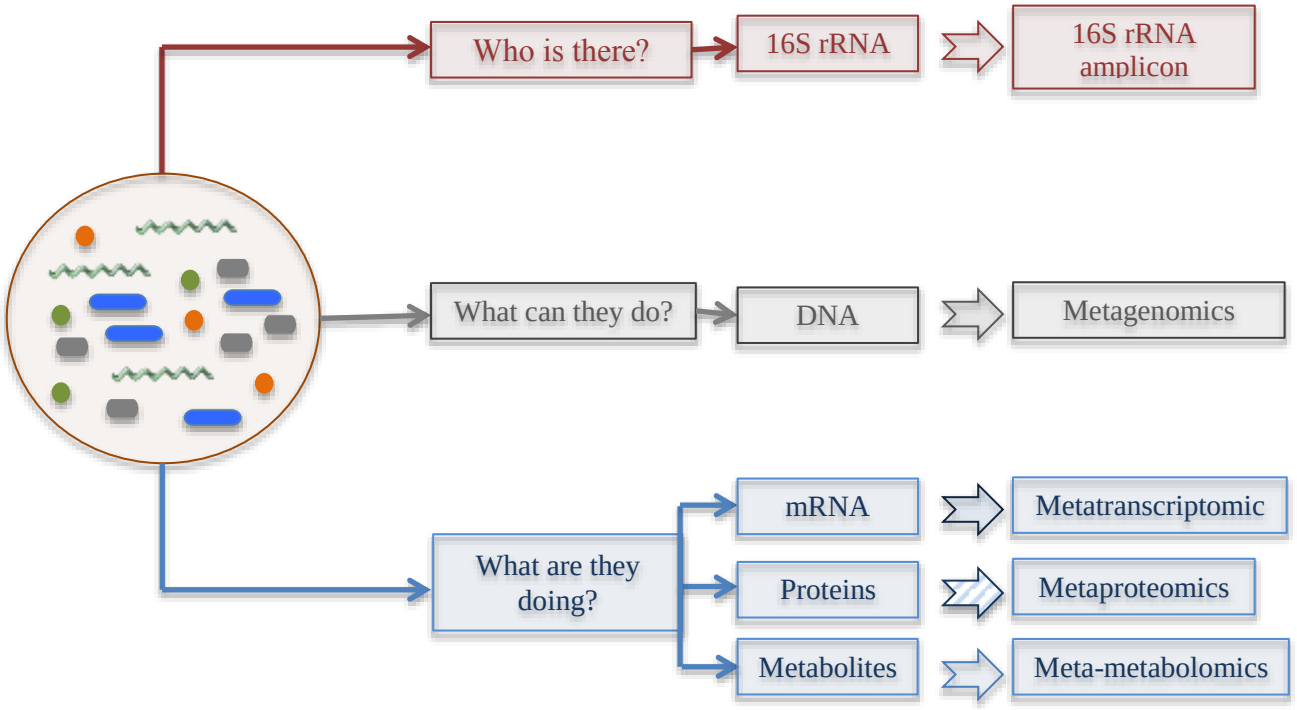
2013), reflecting its low cost and efficiency in decoding bacterial genomes in hours or days instead of months or years.

With the increased use and high demand for technological improvement, a new technology was recently developed and described as the third-generation sequencing. This approach brought remarkable innovation to the field and no longer requires the DNA amplification step prior to sequencing, leading to a more efficient approach, avoiding potential bias and error generated during the amplification procedure, and reduced labor time compared to that typically encountered in the second-generation approach (Liu et al., 2012). Two technologies are currently available, the single-molecule real-time analysis that is based on fluorescence and, the Nanopore sequencing that is based on electric current (Liu et al., 2012). Both technologies have the ability to capture the data signal in real-time, "during the enzymatic reaction of adding nucleotide in the complementary strand" (Liu et al., 2012).

In conjunction with sequencing-technology, a wide rage of molecular approaches is also available for the detailed taxonomic characterization of microbial communities, as well as their functional potential and metabolic products (Di Bella et al., 2013, Addis et al., 2016). Essentially, the molecular strategy applied to the study of the heterogeneous set of microbial genomes is based on the experimental hypothesis and research goal (**Figure 1**) (Addis et al., 2016). Three approaches based on sequencing technology are currently available: the amplicon sequencing that provides access to the community based on a target gene; metagenomics, in which the entire DNA content of a sample is sequenced; and metatranscriptomics in which the entire mRNA content of a sample is sequenced (Di Bella et al., 2013, Addis et al., 2016). Beyond the molecular approaches that rely on sequencing technology, metaproteomics, which allows characterization of the entire microbial protein complement, as well as metabolomics, which

helps characterize the entire metabolite complement, are now available (Addis et al., 2016). For metaproteomic studies, liquid chromatography separation with high-resolution mass spectrometry are used, and for metabolomics liquid chromatography separation system and proton nuclear magnetic resonance spectroscopy are typically used (Addis et al., 2016).

Figure 1.1: The current available approaches for microbiome research (adapted from Addis et al., 2016).



16S rRNA gene & amplicon sequencing approach

The ribosome, a specialized and complex "molecular machine", is a highly conserved structure among all three domains of life. A ribosome consists of several proteins as well as RNA molecules, and is essential for protein biosynthesis. In prokaryotic cells, the ribosome structure has the sedimentation properties of 70S particles, but when not actively engaged in

protein systhesis, the 70S dissociates into the 50S and 30S particles (Pace et al., 1973). The bacterial ribosome comprises of three distinct RNA molecules, the 23S and 5S rRNAs derived from the 50S subunit, and the 16S rRNA that is derived from the 30S ribosomal subunit (Pace et al., 1973). While the ribosome of bacterial organisms evolved and consequently differentiated in their detailed structure, functional homology has been retained across members of the Bacteria domain (Pace et al., 1973), thus allowing it to be used to study bacterial phylogeny and taxonomy (Woese et al., 1987).

In amplicon sequencing a particular gene is selected, amplified and subsequently sequenced for taxonomic profiling. For such approach, sequencing of an evolutionarily stable marker gene, such as the rRNA genes, is usually used as a target (Rajendhran and Gunasekaran et al., 2011). The 16S rRNA seems to be the most conserved gene among the three rRNA genes of prokaryotes (Woese et al., 1987), due to its impact on the translation fidelity rates (high accuracy of the mRNA sequences being translated into protein sequences) (Moore et al., 2011). Usually if the levels of genomic DNA sequence similarity between bacterial species are 70% or greater, they would have more than 97% 16S rRNA gene sequence identity (Stackebrandt and Goebel et al., 1994). Furthermore, given the potential of the 16S rRNA gene to determine the phylogenetic characteristics and "position" of the prokaryotes, it is, in fact, the housekeeping genetic marker most commonly used for bacterial taxonomy and has guided the reconstruction of the tree of life (Woese et al., 1987, Hug et al., 2016).

The full length of the 16S rRNA marker gene is approximately 1500 base pairs and consists of nine hypervariable regions (V1-V9) ranging from about 30-100 base pairs (Gray et al., 1984). These hypervariable regions are flanked by conserved regions allowing effective discrimination between bacterial taxa (Clarridge et al., 2004, Chakravorty et al., 2007) and

design of universal primers for PCR amplification (Caporaso et al., 2011, Apprill et al., 2015, Yang et al., 2016). Sequencing the full 16S rRNA gene during a microbial community survey, is limited by the sequencing technology and cost (Yang et al., 2016), thus the use of the hypervariable regions for phylogenetic analysis and taxonomic classification has been the approach used by many research groups in microbiome studies (Gill et al., 2006, Dominguez-Bello et al., 2010, Boutin et al., 2015, Lau et al., 2016, Ziegler et al., 2017), including microbiome researchers that focus on clinical science (Hansen and Bedard et al., 2013, Sherry et al., 2013, Park et al., 2016) and prevention (Deurenberg et al., 2017).

Upstream laboratory analysis and its impact on microbial profiling

Rapidly developing sequencing methods are enhancing our ability to understand and define microbiomes and are showing their promise as an applicable diagnostic tool. For such technology, efficient isolation of the genomic DNA from the heterogeneous microbial members of different habitats is crucial for an accurate taxonomic profiling. Members of the Bacterial domain have distinct life styles (e.g. intracellular, extracellular, anaerobic, aerobic), ability of making specialized cells (e.g. endospores), and cell wall morphology (e.g. gram-positive, gram-negative, lack of cell wall), thus a DNA extraction method that can effectively lyse the diverse microbial cells without extensively damaging their genomes is required (Bag et al., 2016). For instance, rigorous treatment during DNA extraction could affect DNA quality, while mild treatment may cause partial lysis particularly for gram-positive bacteria that encompass a thicker layer of peptidoglycan (Bag et al., 2016).

Optimization of cell lysis for cutting edge downstream applications, such as next generation DNA sequencing, has been the focus of several studies (Koskinen et al., 2009,

Quigley et al., 2012, Bag et al., 2016). For instance, recently the International Human Microbiome Standards (IHMS) project recently designed a guideline for standard operating procedures to optimize community DNA extraction methods and data quality from human fecal samples (http://www.microbiome-standards.org). The need for methods, references and standard protocols was also a theme for discussion in the US National Science and Technology Council Committee report, in which it was proposed that "resolving differences in method biases for sequencing would permit honest comparisons between studies" and would improve the reliability of the microbiome data (Stulberg et al., 2016).

Studies have also been evaluating DNA extraction methods from milk samples (Cremonesi et al., 2006, Quigley et al., 2012) due to their great importance to the dairy industry and to human and animal health. The typical nutrient components, such as fat, proteins and calcium, of milk biological fluid are PCR inhibitors and can compromise DNA amplification (Wilson et al., 1997). Previous experiments focused on improving DNA isolation from raw milk and cheese milk (Quigley et al., 2012) as well as from sterile milk inoculated with gram-positive bacteria (e.g. *Staphylococcus* spp.) (Cremonesi et al., 2006). However, milk samples from clinical cases of intra-mammary gland infection (mastitis) encounter extra PCR inhibition factors, such as bacterial and mammalian cellular debris (Wilson et al., 1997), and milk samples from non-clinical cases typically have few bacteria (Zadoks et al., 2014), thus recovering bacterial DNA from these samples might be more challenging.

It is common laboratory practice to isolate the genomic DNA from a milk pellet (Rasolofo et al., 2010, Quigley et al., 2012, Quigley et al., 2013), however this can be an issue regarding accurate metagenomic DNA isolation due to the ability of some bacterial species to bind to milk fat globules and/or to the milk fat globule membrane (Ali-Vehmas et al., 1997, Batt.

and Tortorello et al., 2014). So far, no attempt has been made to identify an appropriate and accurate method for metagenomic DNA extraction from non-clinical and clinical cases of mastitis.

HOST-MICROBIOME INTERACTION FROM THE NEXT-GENERATION SEQUENCING PERSPECTIVE

Microbiome and its host health

Mounting evidence on the microbiome and its impact on host health and metabolic functions has been generated by next-generation sequencing over the last few years. Remarkably, the microbiome is now referred to as the "new biomarker" of health (Shukla et al., 2017) mainly due to its role in maintaining host physiology (Sommer and Bäckhed et al., 2013, Barrett and Wu et al., 2017), maturation and "education" of the immune system (Kelly et al., 2007, Chung et al., 2012) and its potential to mediate host metabolic development (Cho et al., 2012, Cox et al., 2014). Since coevolution of mammals and their microbiota has occurred over millions of years (Ley et al., 2008) is not a surprise that this tight connection between host and microbiota exists.

In veterinary medicine, symbiosis between ruminant host and microbial populations is essential for animal survival and existence, mainly due to the conversion of plant and grain materials, consumed by the host, into available energy resources that are subsequently absorbed and metabolized by the host (Hungate et al., 1966). Given its complexity and importance to the ruminant existence, the bovine gastrointestinal tract microbiota has been extensively investigated (Khafipour et al., 2009, Jami and Mizrahi, 2012, Jami et al., 2013, Meale et al., 2016, Dill-McFarland et al., 2017). However, it is becoming increasingly apparent that microbial communities in other anatomical sites, such as mammary gland (Oikonomou et al., 2012,

Oikonomou et al., 2014, Addis et al., 2016) and airways (Holman et al., 2017), are also relevant for bovine health.

Bovine colostrum microbiome

Until quite recently, the mammary gland as well as its biological fluids were believed to be sterile (Tolle et al., 1980) and the detection of bacterial organisms thought to bea result of external contamination from the environment, cow skin or oral cavity of the progeny (Addis et al., 2016). More currently, Martin et al. (2003) found that bacterial species isolated from the mammary gland were genotypically distinct from those observed in the breast skin habitat of the same individual (Martín et al., 2003). The existence of an "entero-mammary pathway" has been also proposed and described as an endogenous route allowing bacterial organisms present in the intestinal lumen to reach the mammary gland by traveling through the mesenteric lymph nodes (Boehm and Stahl et al., 2007, Fernández et al., 2013, Addis et al., 2016). These insights support the trustworthiness of finding an indigenous microbiota in the biological fluids colostrum and milk that is not merely external contamination.

A comprehensive analysis of temporal changes of the human milk microbiome was recently performed by Cabrera-Rubio et al (2010), and the microbial community of colostrum and milk samples from the first and sixth months after delivery were assessed (Cabrera-Rubio et al., 2012). Colostrum was found to harbor a diverse microbiota that were distinct from that observed in milk samples, and bacteria such as *Weisella*, *Leuconostoc*, *Staphylococcus*, *Streptococcus*, and *Lactococcus* were the most common genera detected (Cabrera-Rubio et al., 2012). To date, only one study had explored the microbial community of bovine colostrum (Lindner et al., 2011). The Lindner et al (2011) experiment, found 29 bacterial strains in bovine

colostrum samples including *Lactococcus casei*, *Staphylococcus pseudintermedius*, *Staphylococcus chromogenes*, *Bifidobacterium pseudolongum*, and *Propionibacterium acnes* species (Lindner et al., 2011). However, the present study was limited in its accuracy in describing the colostrum whole bacterial community due to the molecular analysis used. Here, the 16S rRNA genes were generated from DNA extracted from colostrum bacterial isolates and not from the actual colostrum samples.

Inflammation of the mammary gland parenchyma, so-called bovine mastitis, is a huge economic burden for the dairy industry worldwide and is mainly triggered by bacterial infection (Federation, 1987, Oviedo-Boyso et al., 2007). The estimation of the cost per case of clinical mastitis that occurs in the first 30 days after parturition is approximately $444 (Rollin et al., 2015). Colostrum is the first milk produce by dairy cows and in the case of multiparous cows it usually happen after a non-lactating period called dry period, which is an important period for the replacement of senescent mammary epithelial cells and the elements produced by those cells (Capuco et al., 1997). Several studies have shown a correlation between an increased risk of postpartum mammary gland infections with the occurrence of dry period mammary infections (Aarestrup and Jensen et al., 1997, Green et al., 2002, Bradley and Green, 2004). Furthermore, due to the high numbers of intra-mammary infections in both the dry period and the early fresh period (Natzke et al., 1981), the need for new intervention procedures during these periods has been proposed to improve udder health and milk production (Natzke et al., 1981).

Despite the aforementioned issue, until now the question of whether the bovine colostrum microbiome can explain the intra-mammary gland microbial composition shifts toward health or disease along the lactation has not been addressed. Additionally, the identification of colostrum bacterial profiles as a tool for mastitis diagnosis or prevention remains unexplored.

Bovine respiratory tract microbiome

The upper respiratory tract (URT) hosts a complex microbial community (Margolis et al., 2010) that is comprised of commensal microorganisms and also by potential pathogenic agents. However, the specific makeup of the bovine URT microbiome has not been clearly defined. Most studies of URT microflora to date were either culture- or PCR-based surveys that targeted few bacterial groups and, therefore, are incomplete and potentially biased (Allen, 1991, Allen et al., 1992, Arcangioli et al., 2012). Imbalances of the URT ecosystem in humans can result in overgrowth of and invasion by bacterial pathogens, causing respiratory or invasive diseases, especially in children with immature immune systems (Bosch et al., 2013). Analogously, given the high incidence of bovine respiratory disease (BRD) in both dairy (USDA, 2002, 2007) and beef cattle (Griffin et al., 2010), microbial imbalances in the URT microbiome, particularly during the first weeks of life, may potentially lead to respiratory infection. Furthermore, the anatomical connection between the nasophrynx and both the lower respiratory tract and middle ear has been shown as a potential migratory route of pathogenic agents to the lower respiratory tract and middle ear (Murphy et al., 2009). To date, the characterization of the URT microbiome was only carried out by few studies (Holman et al., 2015a, Holman et al., 2015b, Timsit et al., 2016, Holman et al., 2017).

The bovine URT microbiome was firstly described by Holman et al. (2015). In this study, the bacterial community of nasopharynx from ten feedlot cattle was assessed at the day of feedlot entry and also ≥ 60 days later (Holman et al., 2015a). The most abundant phyla detected in the nasopharynx were Proteobacteria and Firmicutes, and the most abundant families were

Pseudomonadaceae, *Moraxellaceae*, *Enterococcaceae*, *Lachnospiraceae*, and *Alteromonadaceae*. It was also found that bovine nasopharyngeal bacterial community is relatively unstable during the first 60 days at feedlot. Bacterial families such as *Enterobacteriaceae*, *Oxalobacteraceae*, and *Pasteurellaceae* were relatively more abundant at the day of feedlot entrance and *Micrococcaceae* and *Pseudomonadaceae* families were relatively more abundant at the sampling time point ≥ 60 days after feedlot entrance (Holman et al., 2015a).

In a more recently published study conducted by the same group, nasopharyngeal microbiota of beef heifers (n=14) from a closed and disease-free herd before and after being transported to the feedlot (Holman et al., 2017) were assessed and also indicated temporal changes in the nasopharyngeal bacterial community. Sample were taken prior to the transport, at the day they arrived at the feedlot, and three samples after feedlot placement on days 2, 7, and 14. The bovine nasopharyngeal habitat harbored a core microbiota composed of 52 shared Operational Taxonomic Units (OTUs) among all time points evaluated. *Mycoplasma*, *Psychrobacter*, *Amnibacterium*, as well as *Acinetobacter* were the most abundant genera detected in the core nasopharyngeal microbiome of the beef heifers (Holman et al., 2017).

In regards to the effect of the microbiome on animal health, Holman et al. (2015a) found that the nasopharynx of animals diagnosed with BRD presented a less diverse microbiota and a reduced number of OTUs than that observed in animals that remained healthy during feedlot placement. Additionally, the *Lactobacillaceae* and *Bacillaceae* families were more abundant at day 0 in cattle that remained healthy compared to animals diagnosed with BRD after feedlot entrance (Holman et al., 2015a). The authors did not observe any bacterial profile associated with BRD a finding that might be explained by the time points used to evaluate the disease-

microbiota associations. Animals were diagnosed with BRD were treated with antibiotic two/or more months prior to the last sample collection (≥ 60 days), thus the microbiota detected in nasopharynx at the sampling time point ≥ 60 after feedlot entrance are not an accurate representation of a "diseased microbiota" and might explain the above-mentioned finding of no bacterial profile associated with BRD.

Despite the noteworthy progress on deciphering the complex bovine respiratory tract microbiome, the URT taxonomic structure and its temporal evolution starting from birth, is still underexplored. Up to the present time, the question the existence of a URT microbial pattern associated with BRD and otitis media specifically during the initial and critical period of life has not been accurately addressed.

HOST MICROBIOME ORIGIN

Maternal microbiome effect on the offspring initial microbial colonization

Early-life microbial colonization has been shown to play a role in the neonatal immune system maturation (Kelly et al., 2007, Chung et al., 2012) and in its metabolic development (Cho et al., 2012, Cox et al., 2014). The first exposure to microbes might happen during fetal development (Romano-Keeler and Weitkamp et al., 2015) due to contamination of the amniotic fluid in intrauterine infection of pregnant women (Goldenberg et al., 2000). More recently, a study performed by Aagaard et al. (2014) detected the presence of a unique microbiome in the placenta of healthy women, in which non-pathogenic commensal microbiota showed to be part of this habitat, indicating that maternal transmission of microorganisms to the fetus in healthy pregnancy might also exist. Additionally, Goldenberg et al. (2000) demostrated culture-positive

chorioamniotic tissue from preterm labor cases presenting intact membranes, which also supports the maternal bacterial transference via placenta (Goldenberg et al., 2000).

Beyond the potential placenta-fetus bacterial transmission route, studies have revealed additional influences on the microbiota acquired by newborns. Using a cohort of maternal-newborn matched subjects, Dominquez-Bello et al. (2010) found that babies born by cesarean section (C-section) harbored a microbiota similar to that observed in maternal skin surfaces while babies born vaginally acquired microbiota resembling their mother's vaginal bacterial community (Dominguez-Bello et al., 2010). The mode of delivery, more specifically C-section, has been also associated with an increased risk of several autoimmune diseases, which was attributed to the disruption of mother-newborn microbial transmission (Mueller et al., 2015a). The interruption of the microbial exchange by C-section has been shown, for instance, to increase the risk of childhood obesity (Huh et al., 2012, Mueller et al., 2015b) and asthma (Metsälä et al., 2015, Mueller et al., 2015b).

After birth, the maternal microbiota continues to influence progeny microbiota through breast feeding (Mueller et al., 2015a). Studies have been shown a diverse microbial community in breast milk (Cabrera-Rubio et al., 2012, Jeurink et al., 2013) and bacteria such as *Bifidobacterium* and *Lactobacillus* spp., were found to be part of breast milk microbiota and to be transferred to the neonate gut (Martín et al., 2007, Solís et al., 2010). Additionally, Kubinak and Round (2016) recently discussed the importance of the maternal antibodies that are passed in milk to progeny and how they influence composition of the progeny gut microbiota composition (Kubinak and Round, 2016). It was proposed that the adaptive immune system not only influences microbial fitness, promoting microbial diversity and mediating symbiosis, but also

functions to reduce the expression of potentially harmful epitopes of commensals and prevent lethal dissemination by controlling mucosa associated communities (Kubinak and Round, 2016).

In veterinary medicine, the origin of the bovine neonatal calf early microbial community and the complex dynamics of their transmission and their site-specific colonization in the neonate remain unexplored.

CONCLUSION AND RESEARCH OBJECTVES

Given the key role of the microbiome on host health, the determination of what constitutes a healthy microbiome and the variability found across populations and through temporal changes are crucial for elucidating the deviations that are associated with disease states. Microbes colonize all anatomical areas of the host, including the gut, airways and mammary gland, however the majority of the bovine microbiome studies are focused on the gut habitat and few have actually explored others anatomical sites. Additionally, given the importance of the microbial DNA isolation procedure on the accuracy of the amplicon sequencing to truly characterize the taxonomic profile of a sample and its sensitivity in diagnose the causative agents of clinical samples, highlight the need of more efficient DNA extraction methods are needed for a more accurate and precise taxonomic description.

The objects of this research were to:

1) Investigate the variation of the microbiota composition in different milk sample types, with emphasis on the community genomic DNA obtained from raw whole milk, milk fat, milk casein-pellet, and milk casein-pellet and fat combined using two commonly used DNA extraction kits.

2) Use high-throughput sequencing of the 16S rRNA gene amplicons to assess the microbial community of colostrum of Holstein dairy cows and evaluate its potential association with early clinical mastitis.
3) Longitudinally characterize the URT microbiome of healthy and unhealthy Holstein dairy calves by using high-throughput sequencing of the 16S rRNA gene.
4) Characterize the vaginal and fecal microbiotas of Holstein dairy cows within the last week of pregnancy and compare them to the fecal and URT microbiotas of their offspring to determine 1) the microbial taxa shared between dam and calf, and 2) the influence of the dam pre-partum vaginal and fecal microbiotas on calf URT and fecal microbial compositions at days 3, 14, and 35 of life.

REFERENCES

Aagaard, K. M. 2014. Author response to comment on "the placenta harbors a unique microbiome". Sci Transl Med 6(254):254lr253.

Aarestrup, F. M. and N. E. Jensen. 1997. Prevalence and duration of intramammary infection in Danish heifers during the peripartum period. J Dairy Sci 80(2):307-312.

Addis, M. F., A. Tanca, S. Uzzau, G. Oikonomou, R. C. Bicalho, and P. Moroni. 2016. The bovine milk microbiota: insights and perspectives from -omics studies. Mol Biosyst.

Ali-Vehmas, T., P. Westphalen, V. Myllys, and M. Sandholm. 1997. Binding of *Staphylococcus aureus* to milk fat globules increases resistance to penicillin-G. J Dairy Res 64(2):253-260.

Allen, J. W., L. Viel, K. G. Bateman, and S. Rosendal. 1992. Changes in the bacterial flora of the upper and lower respiratory tracts and bronchoalveolar lavage differential cell counts in feedlot calves treated for respiratory diseases. Can J Vet Res 56(3):177-183.

Allen, J. W., Viel, L., Bateman, K. G., Rosendal, S., Shewen, P. E., Physick-Sheard, P. 1991. The microbial flora of the respiratory tract in feedlot calves: associations between nasopharyngeal and bronchoalveolar lavage cultures. Can J Vet Res 55(4):341-346.

Apprill, A., S. McNally, R. Parsons, and L. Weber. 2015. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11bacterioplankton. Aquat Microb Ecol 75:129-137.

Arcangioli, M. A., H. Aslan, F. Tardy, F. Poumarat, and D. Le Grand. 2012. The use of pulsed-field gel electrophoresis to investigate the epidemiology of Mycoplasma bovis in French calf feedlots. Vet J 192(1):96-100.

Bag, S., B. Saha, O. Mehta, D. Anbumani, N. Kumar, M. Dayal, A. Pant, P. Kumar, S. Saxena, K. H. Allin, T. Hansen, M. Arumugam, H. Vestergaard, O. Pedersen, V. Pereira, P. Abraham, R. Tripathi, N. Wadhwa, S. Bhatnagar, V. G. Prakash, V. Radha, R. M. Anjana, V. Mohan, K. Takeda, T. Kurakawa, G. B. Nair, and B. Das. 2016. An Improved Method for High Quality Metagenomics DNA Extraction from Human and Environmental Samples. Sci Rep 6:26775.

Barrett, K. E. and G. D. Wu. 2017. Influence of the microbiota on host physiology - moving beyond the gut. J Physiol 595(2):433-435.

Batt C.A. and Tortorello ML. 2014. Encyclopedia of food microbiology. Vol. 1. Second ed. Microbiology of cream and butter. Elsevier.

Boehm, G. and B. Stahl. 2007. Oligosaccharides from milk. J Nutr 137(3 Suppl 2):847S-849S.

Bosch, A. A., G. Biesbroek, K. Trzcinski, E. A. Sanders, and D. Bogaert. 2013. Viral and bacterial interactions in the upper respiratory tract. PLoS Pathog 9(1):e1003057.

Boutin, S., Graeber, S. Y., Weitnauer, M., Panitz, J., Stahl, M., Clausznitzer, D., Kaderali, L., Einarsson, G., Tunney, M. M., Elborn, J. S., Mall, M. A., Dalpke, A. H. 2015. Comparison of microbiomes from different niches of upper and lower airways in children and adolescents with cystic fibrosis. PLoS One 10(1):e0116029.

Bradley, A. J. and M. J. Green. 2004. The importance of the nonlactating period in the epidemiology of intramammary infection and strategies for prevention. Vet Clin North Am Food Anim Pract 20(3):547-568.

Cabrera-Rubio, R., M. C. Collado, K. Laitinen, S. Salminen, E. Isolauri, and A. Mira. 2012. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 96(3):544-551.

Caporaso, J. G., C. L. Lauber, W. A. Walters, D. Berg-Lyons, C. A. Lozupone, P. J. Turnbaugh, N. Fierer, and R. Knight. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci U S A 108 Suppl 1:4516-4522.

Capuco, A. V., R. M. Akers, and J. J. Smith. 1997. Mammary growth in Holstein cows during the dry period: quantification of nucleic acids and histology. J Dairy Sci 80(3):477-487.

Chakravorty, S., D. Helb, M. Burday, N. Connell, and D. Alland. 2007. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. J Microbiol Methods 69(2):330-339.

Cho, I., S. Yamanishi, L. Cox, B. A. Methé, J. Zavadil, K. Li, Z. Gao, D. Mahana, K. Raju, I. Teitler, H. Li, A. V. Alekseyenko, and M. J. Blaser. 2012. Antibiotics in early life alter the murine colonic microbiome and adiposity. Nature 488(7413):621-626.

Chung, H., S. J. Pamp, J. A. Hill, N. K. Surana, S. M. Edelman, E. B. Troy, N. C. Reading, E. J. Villablanca, S. Wang, J. R. Mora, Y. Umesaki, D. Mathis, C. Benoist, D. A. Relman, and D. L. Kasper. 2012. Gut immune maturation depends on colonization with a host-specific microbiota. Cell 149(7):1578-1593.

Clarridge, J. E. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin Microbiol Rev 17(4):840-862, table of contents.

Cox, L. M., S. Yamanishi, J. Sohn, A. V. Alekseyenko, J. M. Leung, I. Cho, S. G. Kim, H. Li, Z. Gao, D. Mahana, J. G. Zárate Rodriguez, A. B. Rogers, N. Robine, P. Loke, and M. J. Blaser. 2014. Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. Cell 158(4):705-721.

Cremonesi, P., B. Castiglioni, G. Malferrari, I. Biunno, C. Vimercati, P. Moroni, S. Morandi, and M. Luzzana. 2006. Technical note: Improved method for rapid DNA extraction of mastitis pathogens directly from milk. J Dairy Sci 89(1):163-169.

Deurenberg, R. H., E. Bathoorn, M. A. Chlebowicz, N. Couto, M. Ferdous, S. García-Cobos, A. M. Kooistra-Smid, E. C. Raangs, S. Rosema, A. C. Veloo, K. Zhou, A. W. Friedrich, and J. W. Rossen. 2017. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol 243:16-24.

Di Bella, J. M., Y. Bao, G. B. Gloor, J. P. Burton, and G. Reid. 2013. High throughput sequencing methods and analysis for microbiome research. J Microbiol Methods 95(3):401-414.

Dill-McFarland, K. A., J. D. Breaker, and G. Suen. 2017. Microbial succession in the gastrointestinal tract of dairy cows from 2 weeks to first lactation. Sci Rep 7:40864.

Dominguez-Bello, M. G., E. K. Costello, M. Contreras, M. Magris, G. Hidalgo, N. Fierer, and R. Knight. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota

across multiple body habitats in newborns. Proc Natl Acad Sci U S A 107(26):11971-11975.

Federation, I. D. 1987. Machine Milking and Mastitis. Bulletin of the International Dairy Federation

Fernández, L., S. Langa, V. Martín, A. Maldonado, E. Jiménez, R. Martín, and J. M. Rodríguez. 2013. The human milk microbiota: origin and potential roles in health and disease. Pharmacol Res 69(1):1-10.

Feughelman, M., R. L angridge , W. E. Seeds, A. R. Stokes, H. R. Wilson, C. W. Hooper, M. H. Wilkins, R. K. Barclay, and L. D. Hamilton. 1955. Molecular structure of deoxyribose nucleic acid and nucleoprotein. Nature 175(4463):834-838.

Gill, S. R., M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. Science 312(5778):1355-1359.

Goldenberg, R. L., J. C. Hauth, and W. W. Andrews. 2000. Intrauterine infection and preterm delivery. N Engl J Med 342(20):1500-1507.

Goodwin, S., J. D. McPherson, and W. R. McCombie. 2016. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet 17(6):333-351.

Gray, M. W., D. Sankoff, and R. J. Cedergren. 1984. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. Nucleic Acids Res 12(14):5837-5852.

Green, M. J., L. E. Green, G. F. Medley, Y. H. Schukken, and A. J. Bradley. 2002. Influence of dry period bacterial intramammary infection on clinical mastitis in dairy cows. J Dairy Sci 85(10):2589-2599.

Griffin, D., M. M. Chengappa, J. Kuszak, and D. S. McVey. 2010. Bacterial pathogens of the bovine respiratory disease complex. Vet Clin North Am Food Anim Pract 26(2):381-394.

Handelsman, J., M. R. Rondon, S. F. Brady, J. Clardy, and R. M. Goodman. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol 5(10):R245-249.

Hansen, A. R. and P. L. Bedard. 2013. Clinical application of high-throughput genomic technologies for treatment selection in breast cancer. Breast Cancer Res 15(5):R97.

Heather, J. M. and B. Chain. 2016. The sequence of sequencers: The history of sequencing DNA. Genomics 107(1):1-8.

Holman, D. B., T. A. McAllister, E. Topp, A. D. Wright, and T. W. Alexander. 2015. The nasopharyngeal microbiota of feedlot cattle that develop bovine respiratory disease. Vet Microbiol 180(1-2):90-95.

Holman, D.B., Timsit, E. and Alexander, T.W., 2015. The nasopharyngeal microbiota of feedlot cattle. Nat. Sci. Rep 5.

Holman, D. B., E. Timsit, S. Amat, D. W. Abbott, A. G. Buret, and T. W. Alexander. 2017. The nasopharyngeal microbiota of beef cattle before and after transport to a feedlot. BMC Microbiol 17(1):70.

Hoy, M. A. 2013. DNA Sequencing and the Evolution of the "-Omics". Pages 251-305 in Insect Molecular Genetics. Third ed. Academic Press.

Hug, L. A., B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, and J. F. Banfield. 2016. A new view of the tree of life. Nat Microbiol 1:16048.

Huh, S. Y., S. L. Rifas-Shiman, C. A. Zera, J. W. Edwards, E. Oken, S. T. Weiss, and M. W. Gillman. 2012. Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study. Arch Dis Child 97(7):610-616.

Hungate, R. 1966. The rumen and its microbes. Academic Press Inc, New York.

Jami, E., A. Israel, A. Kotser, and I. Mizrahi. 2013. Exploring the bovine rumen bacterial community from birth to adulthood. ISME J 7(6):1069-1079.

Jami, E. and I. Mizrahi. 2012. Similarity of the ruminal bacteria across individual lactating cows. Anaerobe 18(3):338-343.

Jeurink, P. V., J. van Bergenhenegouwen, E. Jiménez, L. M. Knippels, L. Fernández, J. Garssen, J. Knol, J. M. Rodríguez, and R. Martín. 2013. Human milk: a source of more life than we imagine. Benef Microbes 4(1):17-30.

Kelly, D., T. King, and R. Aminov. 2007. Importance of microbial colonization of the gut in early life to the development of immunity. Mutat Res 622(1-2):58-69.

Khafipour, E., S. Li, J. C. Plaizier, and D. O. Krause. 2009. Rumen microbiome composition determined using two nutritional models of subacute ruminal acidosis. Appl Environ Microbiol 75(22):7115-7124.

Koskinen, M. T., J. Holopainen, S. Pyörälä, P. Bredbacka, A. Pitkälä, H. W. Barkema, R. Bexiga, J. Roberson, L. Sølverød, R. Piccinini, D. Kelton, H. Lehmusto, S. Niskala, and L. Salmikivi. 2009. Analytical specificity and sensitivity of a real-time polymerase chain reaction assay for identification of bovine mastitis pathogens. J Dairy Sci 92(3):952-959.

Kubinak, J. L. and J. L. Round. 2016. Do antibodies select a healthy microbiota? Nat Rev Immunol 16(12):767-774.

Lau, J. T., F. J. Whelan, I. Herath, C. H. Lee, S. M. Collins, P. Bercik, and M. G. Surette. 2016. Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. Genome Med 8(1):72.

Ley, R. E., C. A. Lozupone, M. Hamady, R. Knight, and J. I. Gordon. 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol 6(10):776-788.

Lindner, J. D. D., Santarelli, M., Yamaguishi, C. T., Soccol, C. R., & Neviani, E. 2011. Recovery and identification of bovine colostrum microflora using traditional and molecular approaches. Page 364. Food Technology and Biotechnology.

Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, and M. Law. 2012. Comparison of next-generation sequencing systems. J Biomed Biotechnol 2012:251364.

Loman, N. J., R. V. Misra, T. J. Dallman, C. Constantinidou, S. E. Gharbia, J. Wain, and M. J. Pallen. 2012. Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 30(5):434-439.

Margolis, E., A. Yates, and B. R. Levin. 2010. The ecology of nasal colonization of Streptococcus pneumoniae, Haemophilus influenzae and Staphylococcus aureus: the role of competition and interactions with host's immune response. BMC Microbiol 10:59.

Martín, R., G. H. Heilig, E. G. Zoetendal, H. Smidt, and J. M. Rodríguez. 2007. Diversity of the Lactobacillus group in breast milk and vagina of healthy women and potential role in the colonization of the infant gut. J Appl Microbiol 103(6):2638-2644.

Martín, R., S. Langa, C. Reviriego, E. Jimínez, M. L. Marín, J. Xaus, L. Fernández, and J. M. Rodríguez. 2003. Human milk is a source of lactic acid bacteria for the infant gut. J Pediatr 143(6):754-758.

Meale, S. J., S. Li, P. Azevedo, H. Derakhshani, J. C. Plaizier, E. Khafipour, and M. A. Steele. 2016. Development of Ruminal and Fecal Microbiomes Are Affected by Weaning But Not Weaning Strategy in Dairy Calves. Front Microbiol 7:582.

Metsälä, J., A. Lundqvist, L. J. Virta, M. Kaila, M. Gissler, and S. M. Virtanen. 2015. Prenatal and post-natal exposure to antibiotics and risk of asthma in childhood. Clin Exp Allergy 45(1):137-145.

Metzker, M. L. 2010. Sequencing technologies - the next generation. Nat Rev Genet 11(1):31-46.

Moore, P.B. and Steitz, T.A., 2011. The roles of RNA in the synthesis of protein. Cold Spring Harbor perspectives in biology, 3(11), p.a003780.

Mueller, N. T., E. Bakacs, J. Combellick, Z. Grigoryan, and M. G. Dominguez-Bello. 2015a. The infant microbiome development: mom matters. Trends Mol Med 21(2):109-117.

Mueller, N. T., R. Whyatt, L. Hoepner, S. Oberfield, M. G. Dominguez-Bello, E. M. Widen, A. Hassoun, F. Perera, and A. Rundle. 2015b. Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. Int J Obes (Lond) 39(4):665-670.

Murphy, T. F., L. O. Bakaletz, and P. R. Smeesters. 2009. Microbial interactions in the respiratory tract. Pediatr Infect Dis J 28(10 Suppl):S121-126.

Natzke, R. P. 1981. Elements of mastitis control. J Dairy Sci 64(6):1431-1442.

Oikonomou, G., M. L. Bicalho, E. Meira, R. E. Rossi, C. Foditsch, V. S. Machado, A. G. Teixeira, C. Santisteban, Y. H. Schukken, and R. C. Bicalho. 2014. Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. PLoS One 9(1):e85904.

Oikonomou, G., V. S. Machado, C. Santisteban, Y. H. Schukken, and R. C. Bicalho. 2012. Microbial diversity of bovine mastitic milk as described by pyrosequencing of metagenomic 16s rDNA. PLoS One 7(10):e47671.

Oviedo-Boyso, J., J. J. Valdez-Alarcón, M. Cajero-Juárez, A. Ochoa-Zarzosa, J. E. López-Meza, A. Bravo-Patiño, and V. M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J Infect 54(4):399-409.

Pace, N. R. 1973. Structure and synthesis of the ribosomal ribonucleic acid of prokaryotes. Bacteriol Rev 37(4):562-603.

Park, S. J., M. Saito-Adachi, Y. Komiyama, and K. Nakai. 2016. Advances, practice, and clinical perspectives in high-throughput sequencing. Oral Dis 22(5):353-364.

Quigley, L., R. McCarthy, O. O'Sullivan, T. P. Beresford, G. F. Fitzgerald, R. P. Ross, C. Stanton, and P. D. Cotter. 2013. The microbial content of raw and pasteurized cow milk as determined by molecular approaches. Journal of Dairy Science 96(8):4928-4937.

Quigley, L., O. O'Sullivan, T. P. Beresford, R. Paul Ross, G. F. Fitzgerald, and P. D. Cotter. 2012. A comparison of methods used to extract bacterial DNA from raw milk and raw milk cheese. J Appl Microbiol 113(1):96-105.

Rajendhran, J. and P. Gunasekaran. 2011. Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. Microbiol Res 166(2):99-110.

Rasolofo, E. A., D. St-Gelais, G. LaPointe, and D. Roy. 2010. Molecular analysis of bacterial population structure and dynamics during cold storage of untreated and treated milk. Int J Food Microbiol 138(1-2):108-118.

Rollin, E., K. C. Dhuyvetter, and M. W. Overton. 2015. The cost of clinical mastitis in the first 30 days of lactation: An economic modeling tool. Prev Vet Med 122(3):257-264.

Romano-Keeler, J. and J. H. Weitkamp. 2015. Maternal influences on fetal microbial colonization and immune development. Pediatr Res 77(1-2):189-195.

Sanger, F., G. M. Air, B. G. Barrell, N. L. Brown, A. R. Coulson, C. A. Fiddes, C. A. Hutchison, P. M. Slocombe, and M. Smith. 1977. Nucleotide sequence of bacteriophage phi X174 DNA. Nature 265(5596):687-695.

Sherry, N. L., J. L. Porter, T. Seemann, A. Watkins, T. P. Stinear, and B. P. Howden. 2013. Outbreak investigation using high-throughput genome sequencing within a diagnostic microbiology laboratory. J Clin Microbiol 51(5):1396-1401.

Shukla, S. D., K. F. Budden, R. Neal, and P. M. Hansbro. 2017. Microbiome effects on immunity, health and disease in the lung. Clin Transl Immunology 6(3):e133.

Solís, G., C. G. de Los Reyes-Gavilan, N. Fernández, A. Margolles, and M. Gueimonde. 2010. Establishment and development of lactic acid bacteria and bifidobacteria microbiota in breast-milk and the infant gut. Anaerobe 16(3):307-310.

Sommer, F. and F. Bäckhed. 2013. The gut microbiota--masters of host development and physiology. Nat Rev Microbiol 11(4):227-238.

Stackebrandt, E. and M. Goebel. 1994. Taxonomic Note: A Place for DNA-DNA Reassociation and 16s rRNA Sequence Analysis in the Present Species Definition in Bacteriology. Pages 846-849. Vol. 44, International journal of systematic bacteriology.

Stulberg, E., D. Fravel, L. M. Proctor, D. M. Murray, J. LoTempio, L. Chrisey, J. Garland, K. Goodwin, J. Graber, M. C. Harris, S. Jackson, M. Mishkind, D. M. Porterfield, and A. Records. 2016. An assessment of US microbiome research. Nat Microbiol 1:15015.

Timsit, E., Workentine, M., Schryvers, A.B., Holman, D.B., van der Meer, F. and Alexander, T.W., 2016. Evolution of the nasopharyngeal microbiota of beef cattle from weaning to 40days after arrival at a feedlot. Vet Microbiol 187:75-81.

Tolle, A. 1980. The Microflora of the Udder. Int. Dairy Fed. Bull.:4.

USDA. 2002. Part I: Reference of dairy health and management practices in the United States. in FortCollins (CO): USDA: APHIS: VS, CEAH, National Animal Health Monitoring System. #N377.1202 ed.

USDA. 2007. Heifer Calf Health and Management Practices on U.S. Agriculture, ed. USDA:APHIS:VS:CEAH NRRC Building B, M.S. 2E7 2150 Centre Avenue Fort Collins, CO 80526-811.

Watson, J. D. and F. H. Crick. 1953. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171(4356):737-738.

Website: http://www.microbiome-standards.org. Assessed in 05/18/2017.

Wilkins, M. H. 1956. Physical studies of the molecular structure of deoxyribose nucleic acid and nucleoprotein. Cold Spring Harb Symp Quant Biol 21:75-90.

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63(10):3741-3751.

Woese, C. R. 1987. Bacterial evolution. Microbiol Rev 51(2):221-271.

Yang, B., Y. Wang, and P. Y. Qian. 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. BMC Bioinformatics 17:135.

Zadoks, R. N., R. Tassi, E. Martin, J. Holopainen, S. McCallum, J. Gibbons, and K. T. Ballingall. 2014. Comparison of bacteriological culture and PCR for detection of bacteria in ovine milk--sheep are not small cows. J Dairy Sci 97(10):6326-6333.

Zhou X. and Li Y. 2015. Techniques for Oral Microbiology. Pages 15-40 in Atlas of Oral Microbiology: From Healthy Microflora to Disease. Academic Press.

Ziegler, M., F. O. Seneca, L. K. Yum, S. R. Palumbi, and C. R. Voolstra. 2017. Bacterial community dynamics are linked to patterns of coral heat tolerance. Nat Commun 8:14213.

CHAPTER TWO

EVALUATION OF THE MILK SAMPLE TYPE AND DNA EXTRACTION METHOD FOR CHARACTERIZATION OF MICROBIAL PROFILES OF MASTITIC AND NON-MASTITIC MILK SAMPLES

Svetlana F. Lima*, Marcela L. S. Bicalho*, Ana Carolina Campos*, Leonardo Bringhenti*, Rodrigo C. Bicalho*

*Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, New York, United States of America

ABSTRACT

The advent of amplicon sequencing has revolutionized the study of bacterial communities within a variety of ecosystems, including its application in the clinical sciences as a sensitive diagnostic tool. Therefore, it is of great importance that the method used to extract DNA yields an accurate representation of the complex microbiota of interest. In particular, milk is a challenging sample for DNA extraction because of its physical and chemical characteristics as well as the microbial and host cellular debris that can inhibit PCR amplification. However, there is no established best practice for isolation of microbial DNA from clinical and non-clinical cases of bovine mastitis. In the present study, DNA was extracted from whole milk, milk fat, casein-pellet, and fat plus pellet combined by using two commercial DNA extraction kits [PowerSoil and PowerFood (MoBio/Qiagen Ltd)], giving eight DNA isolation protocols for comparison. Moreover, milk samples from both healthy cows and cows with mastitis caused by infection by *Escherichia coli*, *Klebsiella* spp., and *Streptococcus* spp. were analyzed. The DNA isolation methods were statistically evaluated according to the following criteria: DNA concentration, protocol agreement, microbial representativeness, and reproducibility. Regardless of the milk sample type (i.e., whole milk, fat, pellet, fat + pellet) and DNA extraction kit used, the DNA and amplicon (16S RNA gene) concentrations obtained from mastitic milk samples were sufficient and suitable for next-generation sequencing. However, 47% of the non-mastitic milk samples failed to yield a detectable PCR product. The level of agreement between the two DNA extraction kits with respect to the number of bacterial families detected in non-mastitic milk samples was of 60%. Twenty-eight shared operational taxonomic units comprised the core microbiome identified among all four milk sample types. Furthermore, all eight DNA isolation procedures yielded nucleic acid sufficient for diagnosis of the mastitis causative agent

(*Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp.) and agreed as to the most abundant milk bacterial families. In conclusion, the choice of extraction kit had little effect on sequencing quality and coverage across all four milk sample types. Both PowerSoil and PowerFood showed good reproducibility and were appropriate for next-generation sequencing to identify the main clinical mastitis causative agents. Based on these results, the choice of DNA extraction kit can be guided largely by other factors, including extraction costs and extraction time. In these respects, studies focused on determining the microbial community present in non-mastitic and mastitic milk may benefit from using the PowerFood kit and whole milk.

INTRODUCTION

Inflammation of the mammary gland, also known as mastitis, is arguably the most important disease affecting dairy herds worldwide (Hagnestam-Nielsen and Ostergaard, 2009). Mastitis is a complex disorder mainly triggered by bacterial infection (Federation, 1987, Oviedo-Boyso et al., 2007), typically by coagulase-negative staphylococci, *Bacillus* spp., *Streptococcus* spp., *Staphylococcus aureus,* and *Escherichia coli* (National Mastitis Council, 1999, Oliveira et al., 2013, Levison et al., 2016). Due to its multifactorial etiology and the risk of antibiotic resistance, the best approach to mastitis treatment is to accurately identify the causative agent (Pyörälä, 2009), which typically has been carried out by microbiological culture (National Mastitis Council, 1998, 1999), a standard diagnostic tool in veterinary medicine (Lago et al., 2011). However, because cultures of mastitic milk samples may not always result in bacterial growth, an increasing number of studies has shown the potential of molecular techniques to improve the diagnosis of mastitis, with high sensitivity and specificity (Phuektes et al., 2003, Koskinen et al., 2009, Taponen et al., 2009, Oultram et al., 2017).

Accordingly, for any PCR-based approach, generating high-quality DNA is both critical and a challenge for accurate taxonomic profiling. Milk, in particular, is a challenging sample due to its physical and chemical characteristics, especially its fat, protein and, calcium constituents that act as PCR inhibitors (Wilson, 1997). Furthermore, clinical samples from an infected mammary gland contain additional PCR inhibitory factors such as bacterial and mammalian cellular debris (Wilson, 1997), whereas non-clinical milk samples typically have low bacterial loads (Zadoks et al., 2014). Likewise, different bacterial species may possess distinct cell-structural characteristics (e.g. gram negative, gram positive, "free of cell wall") that may affect DNA recovery, thus the treatment applied to the sample could bias the results of the downstream analysis and consequently the taxonomic profiling.

Milk is a complex biological fluid mainly composed of fat globules and casein micelles (the primary group of milk proteins containing 80% of the total milk protein). All other proteins found in suspension in the fluid phase after precipitation of caseins are grouped together under the name of whey proteins (Dalgleish and Corredig, 2012). In addition to the soluble non-casein proteins, the whey supernatant (milk serum) also contains water and lactose, and the two main components of the serum proteins in bovine milk are α lactalbumin and β lactoglobulin (Walstra, 1999). Typically, isolation of DNA from milk samples is performed by pelleting the casein (Rasolofo et al., 2010, Quigley et al., 2012, Quigley et al., 2013). Recently, Quigley et al. (2012) evaluated seven DNA isolation methods for raw milk and its derivate in terms of their relative success based on DNA yield and purity, as well as the quality of the template for downstream PCR. In the same study, DNA was isolated by resuspending the casein-pellet, which was submitted to different enzymatic and mechanical cell-lysis protocols (Quigley et al., 2012). However, some bacterial species have a diversified lifecycle in the milk environment, and the

growth, location, and distribution of bacterial colonies in dairy products are important factors for the dairy food industry. For instance, starter, non-starter, spoilage, and pathogenic bacteria all become entrapped in the developing casein matrix of dairy foods (Hickey et al., 2015). On the other hand, recent studies have proposed an optimized milk template preparation for more efficient detection of *Mycobacterium avium* subsp. Paratuberculosis by PCR, which involves combining the cream and pellet to produce a milk sample-template with increased PCR sensitivity (Herthnek et al., 2008, Beaver et al., 2016). Furthermore, *Staphylococci* bacteria appear to bind to the fat globules and/or to the milk fat globule membrane, resulting in their increased resistance to antibiotic treatment (Ali-Vehmas et al., 1997, Batt C.A. and Tortorello ML., 2014); thus, PCR-based detection of such bacterial species using traditional DNA isolation methodologies might be reduced.

Here, to identify the best DNA-isolation protocol that accurately isolate the majority of categories of bacterial species (e.g. gram negative, gram positive, bacteria that lack a cell wall) present in mastitic and non-mastitic milk samples, we compared two commonly used DNA extraction kits (PowerFood and PowerSoil/MoBio), which combine physical (heat), mechanical (bead-beating) and chemical cell lysis approaches with a silica-based chromatographic DNA-recovery step for precise bacterial profile determination (Bag et al., 2016). Because, in milk-DNA optimization studies, DNA concentration and PCR efficiency outcomes alone may not serve as definitive indicators of sequence data quality (Weber et al., 2017), we first investigated variation of the microbiota composition in different milk sample types, with emphasis on the DNA obtained from raw whole milk, milk fat, casein-pellet, and casein-pellet and fat combined. Secondly, we evaluated the DNA isolation method that more accurately diagnoses clinical mastitis caused by *Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp., as those are known

to play an important role in bacterial mammary-gland infection in U.S. dairies (USDA, 2014). The DNA isolation methods were statistically evaluated according to the following criteria: DNA concentration, protocol agreement, microbial representativeness, and reproducibility.

MATERIAL AND METHODS

Ethics statement

The research protocol was reviewed and approved by the Cornell University Institutional Animal Care and Use Committee (protocol number 2013-0056). The methods were carried out in accordance with the approved guidelines. A total of 60 cows were sampled.

Ethics statement

The research protocol was reviewed and approved by the Cornell University Institutional Animal Care and Use Committee (protocol number 2013-0056). The methods were carried out in accordance with the approved guidelines. A total of 60 cows was sampled.

Case definition

This study was conducted in one large commercial dairy farm situated in upstate New York due to its long-standing relationship with the Ambulatory Clinic at Cornell University.

Clinical mastitis

Clinical mastitis examination was performed at the milking parlor by farm employees with extensive training and experience and then reconfirmed by one of our research group. Clinical mastitis was defined as the presence of visually abnormal milk (i.e. presence of flakes,

clots, or serous milk) independently of systemic illness and signs of inflammation of the mammary gland during fore-stripping performed at the milking parlor.

Healthy cows

Cows with no visible changes of the secretion and/or the consistency of the mammary tissue were classified as healthy. Additionally, cows were not eligible to be included in the study if they were diagnosed with clinical mastitis in the current lactation and when antimicrobial or anti-inflammatory treatment occurred within the previous 30 d, when cows were within 5 d post-calving, within 30 d of drying off, with visible signs of teat damage, or experiencing concurrent disease.

Sample collection

For culture as well as pre- and post-sequencing outcomes analysis, milk samples were aseptically collected at the milk parlor from a convenience sample of 60 Holstein dairy cows, in which 38 cows were diagnosed with clinical mastitis and 22 were healthy cows. Sampling methods followed standard recommendations by the National Mastitis Council's Laboratory Handbook on Bovine Mastitis (National Mastitis Council, 1999). Briefly, the first streams of milk from each quarter were discarded for mammary gland stimulation, and subsequently the teats were dipped in iodine tincture. Then teats were cleaned and disinfected using 70% ethanol, the first three streams were discarded, and the milk samples were collected into sterile plastic tubes without preservative (Corning Life Sciences, Tewksbury, MA). Approximately 50 ml of milk was collected in a single sterile 50-ml centrifuge tube (Fisher Scientific, Pittsburgh, PA) from each study cow. Milk samples from cows with mastitis were collected from the mastitic

quarter, and milk samples from healthy cows were collected at random from one of the cow's hind quarters. Samples were kept on ice until transported to the laboratory, a 2-ml aliquot was separated for culture analysis and the remaining 48-ml sample was stored at -20 °C for further processing.

Microbiological culture for pathogen identification or confirmation of pathogen-free milk

To identify suitable samples for downstream analysis that are representative of the bacterial community present in each milk sample, we used an on-farm culture system for identification of the main milk pathogens associated with clinical mastitis or confirmation of pathogen-free milk. Milk samples from all cows used in this study were submitted to our laboratory at Cornell University, Ithaca, NY, for bacterial identification using a chromogenic culture system (Accumast®, FERA Animal Health LCC, Ithaca, NY). Using this approach, the expected microbial relative abundances from our study samples, which were further determined by 16S rRNA amplicon sequencing, would be directly associated to the presence of the predicted main pathogen. The choice for this on-farm culture system was due to its proven suitability for use under field conditions and its substantial overall accuracy for detection of common mastitis pathogens, which was previously confirmed by 16S rRNA gene sequencing (Ganda et al., 2016a).

Milk samples were plated on the surface of each selective growth medium (tri-plate system) (Figure 2.1A) using sterile cotton swabs. Plates were aerobically incubated at 37°C for 24 h and subsequently read by one of the research team members. Identification of milk pathogens was performed following instructions of a flowchart developed based on characteristics of growth of American Type Culture Collection (ATCC**)** strains described in Figure 2.1A. Samples were considered mixed infections when two or more clearly distinct

bacterial types in a well-distributed growth pattern were detected, and pathogens were reported. Samples were considered negative when no aerobic bacterial growth was observed in the first 24 h of incubation, following guidelines for accredited diagnostic laboratories.

Description of the DNA extraction kits

PowerFood and PowerSoil Microbial DNA Isolation kits (MoBio/Qiagen Ltd., Germantown, MD) employ physical (heat), mechanical (bead beating) and chemical lysis for proper metagenomic DNA isolation, and the silica column provides a purer metagenomic template (Bag et al., 2016).

PowerFood was designed for "isolation of DNA from tough, food cultured microorganisms" and contains 0.15-mm garnet beads (PowerFood® DNA Isolation Kit, MoBio, protocol). Microbial DNA isolation using the PowerFood kit is based on 6 steps: sample preparation (step 1: centrifugation of 1.8 ml of liquid food or homogenized 0.25 g of food in 0.75 ml of PBS, and then removal of the food residuals), collection of cells (step 2: bacterial cell resuspension with Solution PF1), cell lysis (step 3: samples are exposed to 65°C for 10 minutes followed by a bead beating process for 15 minutes, as suggested by the manufacturer), inhibitor removal (step 4: use of Solution PF2. It removes additional non-DNA organic and inorganic material, including cell debris, and proteins), binding of DNA (step 5: use of Solution PF3. It contains a highly concentrated salt solution allowing binding of DNA), wash procedure (step 4: use of Solution PF4 and PF5. PF4 removes residual salt and other contaminants while allowing the DNA to stay bound to the silica membrane. PF5 ensures complete removal of Solution PF4, which will result in higher DNA purity and yield), lastly an elution process (step 6: use of

Solution PF6 for complete release of the DNA from the silica Spin Filter membrane). As a result, a final volume of 100 µl of the isolated DNA is collected in the final elution step.

PowerSoil was designed for "isolation of environmental samples, including difficult soil types such as compost, sediment and manure" and contains 0.7-mm garnet beads (PowerSoil® DNA Isolation Kit, MoBio, protocol). Microbial DNA isolation using this kit is based on 6 steps, similar to the steps performed in the PowerFood kit. However, in the cell lysis step, samples are exposed to a higher temperature of 70°C for 10 minutes followed by a bead beating process for 15 minutes, as suggested by the manufacturer. Additionally, in the wash step, samples are subjected to a more "intense" clean-up process to improve the purity of the final DNA template. A final volume of 100 µl of the isolated DNA is collected in the final elution step.

Reagent controls were tested for all DNA extraction kits used and subjected to PCR assay to prevent potential contamination attributed to reagent impurity. Negative results for the controls were obtained according to agarose gel analysis. The DNA concentration of milk samples and control samples was evaluated by optical density using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 260 and 280 nm.

Description of the milk sample types

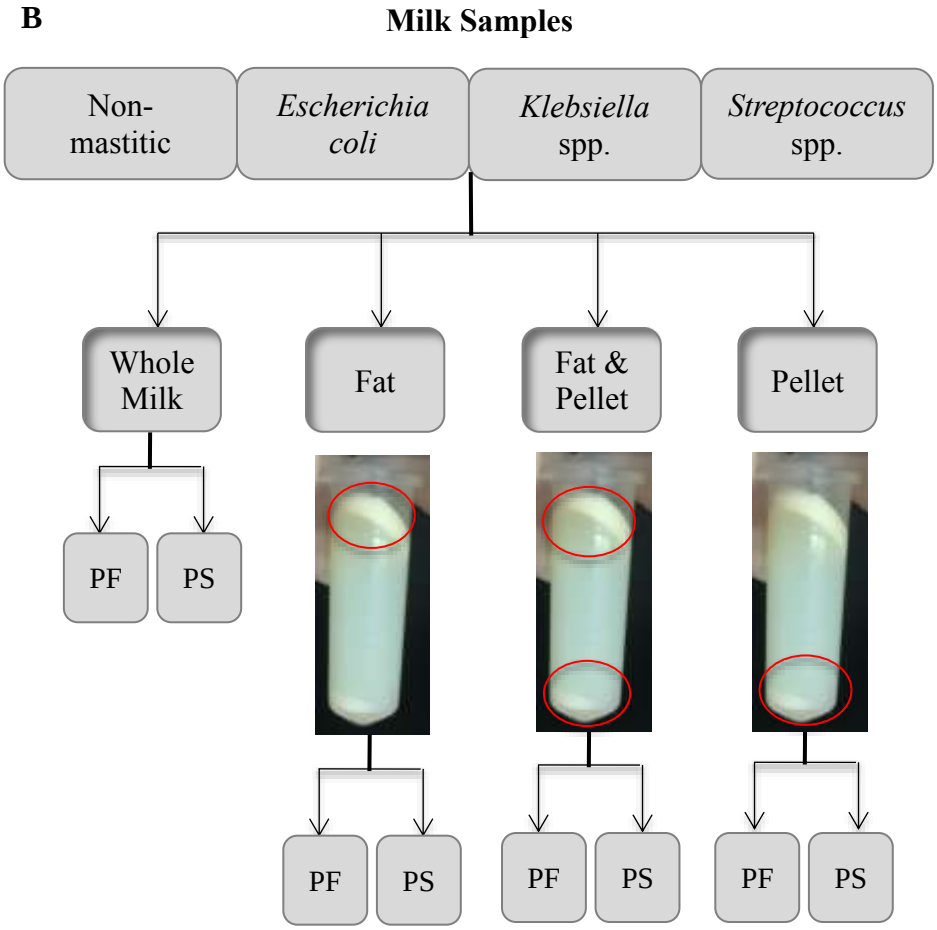
DNA was isolated from four distinct milk sample types: whole milk, fat only, pellet only, and fat and pellet combined, as shown in Figure 2.1B. Frozen samples were thawed on the day of DNA extraction, homogenized and aliquoted in 8 Falcon tubes (15-ml volume) (Fisher Scientific, Pittsburgh, PA) each containing 6 ml of sample. The samples were then subjected to

two different microbial DNA extraction protocols (PowerFood and PowerSoil) for comprehensive microbial community analysis based on 16S rRNA gene sequencing. Samples from the same cow were processed at once, and then subjected to the same PCR assay, purification assay and sequencing batch. Only one member of our research team executed the DNA laboratory procedures, thereby circumventing inter-operator variability.

Figure 2.1: Visual assessment of Escherichia coli, Klebsiella spp. and Streptococcus spp. growth on Accumast plates performed in the laboratory (A). Overview of the experimental design (B). PF, PowerFood microbial DNA isolation kit; PS, PowerSoil microbial DNA isolation kit.

A

B



Metagenomic DNA obtained from whole milk

Two Falcon tubes (15-ml volume) each containing 6 ml of milk were used in this procedure. A total of 250 µl of milk from tube 1 was transferred to a PowerSoil bead tube and a total of 250 µl of milk from tube 2 were mixed with 450 µl of Solution PF1 and transferred to a PowerFood microbead tube. DNA was then extracted according to the manufacturer's instructions.

Metagenomic DNA obtained from milk fat

Two Falcon tubes (15-ml volume) each containing 6 ml of milk were used in this procedure. Milk samples were then transferred to a sterile 2 ml microcentrifuge tube (Eppendorf, Hauppauge, NY) and centrifuged at 16,000 × g for 5 minutes at room temperature. This process was repeated 3 times.

A total of 250 mg of fat content from tube 1 was transferred to a PowerSoil bead tube and DNA extraction was performed according to the manufacturer's instructions. A total of 250 mg of fat content from tube 2 was resuspended in 450 µl of a strong lysing reagent (Solution PF1) from the PowerFood kit, further transferred to a PowerFood microbead tube, followed by DNA extraction according to the manufacturer's instructions.

Metagenomic DNA obtained from milk casein-pellet

Two Falcon tubes (15-ml volume) each containing 6 ml of milk were used in this procedure. Milk samples were then transferred to two 2 ml microcentrifuge tube, centrifuged at 16,000 × g for 5 minutes at room temperature and the supernatant composed of fat and whey was discarded from tube 1 (this process was repeated 3 times). The remaining pellet in tube 1 was

resuspended in 250 µl of the buffer solution used in the PowerSoil bead tubes and transferred to a PowerSoil bead tube. Similarly, supernatant composed of fat and whey was discarded from tube 2 (this process was repeated 3 times) and the remaining pellet was resuspended in 450 µl of Solution PF1 from the PowerFood kit and then transferred to a PowerFood microbead tube. DNA was then extracted according to the manufacturer's instructions.

Metagenomic DNA obtained from milk fat + casein-pellet combined

Two Falcon tubes (15-ml volume) each containing 6 ml of milk were used in this procedure. Milk samples were then transferred to two 2 ml microcentrifuge tube and centrifuged at 16,000 × g for 5 minutes at room temperature. After centrifugation, the whey fraction was removed and discarded from tubes 1 and 2 (this procedure was repeated 3 times); fat and pellet contents were homogenized in 0.75 ml of UltraPure™ distilled water, DNAse- and RNAse-free (Invitrogen Life Science Technologies, Grand Island, NY) for 5 minutes using a vortex with a horizontal adapter.

A total of 250 µl of fat + pellet content from tube 1 was transferred to a PowerSoil bead tube and another 250 mg from tube 2 were mixed in a sterile microcentrifuge tube containing 450 µl of Solution PF1 from the PowerFood kit and then transferred to a PowerFood microbead tube. DNA was then extracted according to the manufacturer's instructions.

PCR amplification of the bacterial 16S rRNA gene and amplicon sequencing

Amplification of the V4 hypervariable region from the 16S rRNA gene was performed by PCR using barcoded primers. Primers 515F and 806R were used according to previously described methods and optimized for the Illumina MiSeq platform (Caporaso, 2012). In total,

280 different 12-bp error-correcting Golay barcodes primers were designed based on "The Earth Microbiome Project" (http://www.earthmicrobiome.org/) (Gilbert et al., 2010). Amplicons were generated in triplicate using 12-300 ng of template DNA, 2X EconoTaq® Plus Green Master Mix (Lucigen®, Middleton, WI) and 10 µM of each primer. The thermocycler conditions included an initial denaturing step of 94°C for 3 min followed by 35 cycles of 94°C for 45 s, 50°C for 1 min and 72°C for 90 s and a final elongation step of 72°C for 10 min. Replicate amplicons were pooled and purified using a Gel PCR DNA Fragment Extraction kit (IBI Scientific, Peosta, IA) and visualized by electrophoresis through 1.2% (wt/vol) agarose gels stained with 0.5 mg/ml ethidium bromide.

Samples that failed to be detected by PCR assay (no bands on the agarose gel) were re-tested; thus, a new PCR and agarose gel procedure were performed for confirmation of the negative result or for recovering of the false-negative sample. The amplicon concentration was evaluated by optical density using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 230, 260 and 280 nm.

Aliquots of milk amplicon samples were diluted to the same concentration and then pooled into one unique run according to individual barcode primers for the 16S rRNA gene, V4 hypervariable region. Final equimolar libraries were sequenced using the MiSeq reagent kit v2 (300 cycles) on the MiSeq platform (Illumina Inc., San Diego, CA).

Bioinformatics

The 16S rRNA gene sequences generated were processed through the open source software pipeline Quantitative Insights Into Microbial Ecology 2 (QIIME 2) version 2017.2 (hpp://qiime2.org). Sequences were demultiplexed using the "demux emp-single" command.

Quality control was performed using DADA2 (Callahan et al., 2016). The DADA2 procedure removes any remaining phiX reads, chimeric sequences and low-quality regions of the sequences. Herein, high-quality bases equal to Q30 (probability of an incorrect base call is 1 in 1000 and the inferred base call accuracy is 99.9%) were observed around position 150 bases, thus sequences were truncated at 150 bases. Sequences were binned into an Operational Taxonomic Unit (**OTU**) based on 100% identity. The training feature "q2-feature-classifier" command using the Greengenes reference database (McDonald, 2012) was created to classify representative sequences from our dataset. The output of this workflow is a classification of reads at multiple taxonomic levels: kingdom (k), phylum (p), class (c), order (o), family (f), genus (g) and species (s).

Statistical analysis

To facilitate data analysis and interpretation of the results whole milk, fat, fat + pellet, and pellet were categorized as milk sample types. Additionally, milk-health status groups were categorized according to the main pathogen identified through standard culture methods, namely, *Escherichia coli* group, *Klebsiella* spp. group, and *Streptococcus* spp. group for cows diagnosed with mastitis; the non-mastitic group consisted of cows diagnosed as healthy.

The Pearson Chi-square test was performed using JMP Pro 12 (SAS Institute Inc.) to evaluate whether the proportion of samples that failed to be detected by PCR (negative amplification of the 16S rRNA gene, indicated by no bands on the agarose gel) differed between milk sample types (whole milk, fat, fat + pellet, and pellet) and within or between DNA extraction kits (PowerFood and PowerSoil).

Differences in the DNA concentration, amplicon concentration, total number of

sequences, and total number of OTUs between milk sample types within or between DNA extraction kits (PowerFood and PowerSoil) were evaluated using the Mixed procedure in SAS 9.4 (SAS Institute Inc.). Bonferroni test was used to adjust for multiple comparisons.

The OTU data obtained from bioinformatics analysis were used to describe the relative abundances of bacterial phyla and families within each milk-health status group across all milk sample types and DNA extraction kits. Each value obtained indicated the percentage relative frequency of reads with 16S rRNA genes annotated to the indicated taxonomic level. The microbiota profile within milk groups was described for the most prevalent phyla and bacterial families using the tabulate function of JMP Pro 12. Graphs representing phyla mean relative abundance were constructed in Excel (Microsoft Corp., Redmond, WA), whereas a Heatmap was generated in JMP Pro 12 to graphically represent the relative distributions of the most common bacterial families found in our samples.

To gain a deeper insight into the dissimilarity levels of milk bacterial communities represented among samples extracted by the eight different DNA isolation methods (4 milk fractions × 2 DNA isolation kits), Venn diagrams (VennDiagram package under RStudio software version 0.99.903; RStudio, Inc) were created for graphical descriptions of the number of unique and shared OTUs at the family level. Tables depicting the unique and shared families, and their respective relative abundances among DNA sample types were generated using the tabulate function of JMP Pro 12.

To test for differential abundance of taxa (at the family level) that might be driven by the eight different DNA isolation methods (4 milk fractions × 2 DNA isolation kits), Kruskal-Wallis tests followed by Benjamini-Hochberg false discovery rate (FDR) calculations were performed using JMP Pro 12. For the mastitic milk groups (*Escherichia coli*, *Klebsiella* spp., *Streptococcus*

spp. groups) the relative abundances of f__Enterobacteriaceae (bacterial family known to comprise the *Escherichia coli* and *Klebsiella* spp. groups) and f__Streptococcaceae (bacterial family known to comprise the *Streptococcus* spp. group), were compared between the DNA isolation methods to determine the efficacy of the procedures in detecting the causative agent of mastitis. The graphs representing family mean relative abundances (MRA) were constructed in Excel and differences with a value of $P \leq 0.05$ were considered significant.

RESULTS

Descriptive statistics

After sample exclusion, due to the following reasons – microbiological culture-positive for non-mastitic milk (n = 5 cows), mastitic milk culture resulted in no-growth (n = 7 cows), and mastitic milk culture resulted in mixed pathogens (n = 7 cows) – milk samples from a total of 41 cows were subjected to DNA extraction followed by PCR screening. Equal volumes of each of the 41 milk samples were aliquoted in 8 microcentrifuge tubes, which were processed as whole milk, fat, fat + pellet, and pellet before being subjected to two different DNA extraction kits (PowerFood or PowerSoil), giving a total of 328 study samples. Of these 328 samples, 120 originated from non-mastitic milk, 40 from mastitic milk due to *Escherichia coli*, 32 from mastitic milk due to *Klebsiella* spp. and 120 samples from mastitic milk due to *Streptococcus* spp. (Table 2.1). Ninety-seven samples were excluded after they failed the DNA detection by PCR (no bands on the agarose gel) (Table 2.1). In total, 231 samples were subjected to amplicon sequencing.

Pre-sequencing outcomes

DNA concentration

The DNA concentration values obtained for mastitic and non-mastitic milk samples across the four milk fractions (whole milk, fat, fat + pellet, and pellet) and two DNA extraction protocols (PowerFood and PowerSoil) are depicted in Table 2.2. On average, whole milk yielded significantly lower DNA concentrations (*P*-value = 0.017; Table 2.2) compared to fat + pellet, but not when compared to the fat or pellet fraction. This difference was mostly attributable to lower performance of the PowerSoil DNA extraction kit (*P*-value = 0.04).

Investigation was also carrying out to determine which DNA extraction Kit provides higher extracted DNA concentration (Table 2.3). Comparison between DNA extraction kits among all four milk-health status groups revealed that samples extracted by PowerFood had significantly higher concentrations of extracted DNA (*P*-value = 0.007, Table 2.3). This difference between kits was particularly noticeable in the whole milk sample (*P*-value = 0.04, Table 2.3). The PowerFood DNA extraction kit also yielded significantly higher DNA concentrations from milk samples from the *Streptococcus* spp. mastitic group (*P*-value 0.03, Table 2.3). However, DNA concentrations did not differ between PowerFood and PowerSoil for the *Escherichia coli* mastitic group, the *Klebsiella* spp. mastitic group, and the non-mastitic group.

DNA amplification success

The percentages of samples that failed to be detected by PCR, characterized by no visible or appropriate size band in the agarose gel (390 base pairs), are depicted in Table 2.1. Differences in the proportion of samples that failed to be detected by PCR assay were not

observed between and within milk sample types and DNA extraction kits, regardless of milk-health status (*Escherichia coli, Klebsiella* spp., *Streptococcus* spp. and non-mastitic groups; Table 2.1A and B). Milk samples from six cows were unsuccessfully extracted by all eight DNA isolation procedures used here. Three of the milk samples were from healthy cows, one sample was from a clinical case of mastitis due to *Escherichia coli* infection, and two samples from clinical cases of mastitis due to *Streptococcus* spp. infection.

Amplicon concentration

In regard to the effect of the DNA isolation procedures on amplicon concentration, on average, samples from whole milk yielded significantly lower amplicon concentrations compared to samples from fat, fat + pellet, and pellet only, when the comparison was made across all milk-health status groups (P-value < 0.0001, Table 2.2). No differences in amplicon concentrations were observed between the pellet and fat fractions, regardless of DNA extraction kit and milk-health status. On the other hand, analysis performed within milk health-status groups found significantly different amplicon concentrations among milk fractions from mastitic milk due to *Klebsiella* spp. when DNA was extracted with the PowerFood kit (P-value < 0.025, Table 2.2) and from mastitic milk due to *Streptococcus* spp. when DNA was extracted with the PowerSoil kit (P-value < 0.002, Table 2.2).

The DNA extraction kits also yielded significantly different amplicon concentrations (P = 0.01, Table 2.3). Moreover, milk samples from mastitic milk due to *Streptococcus* spp. gave significantly higher amplicon concentrations from whole milk and fat fractions by PowerFood in comparison to PowerSoil (P = 0.01, Table 2.3).

Table 2.1. Percentage of samples undetected (n = 97) by polymerase chain reaction (PCR) for the 16S rRNA gene according to milk-health status groups (non-mastitic, n = 136; mastitis caused by *Escherichia coli*, n = 40; *Klebsiella* spp., n = 32; *Streptococcus* spp., n = 120), milk sample type (whole milk, fat, fat + pellet, and pellet), and DNA extraction kit (PorwerFood and PowerSoil) (A). More detailed data concerning the unsuccessful amplification results are depicted in part B. In total, 328 samples were subjected to PCR.

A

	Negative 16S rRNA gene amplification % (n = number of samples)								
		DNA extraction kit			Milk sample types				
Milk-health status	Total	PF[1]	PS[2]	*P*-value	WM[3]	Fat	F+P[4]	Pellet	*P*-value
All groups combined	29.6 (328)	29.9 (164)	29.6 (164)	0.90	28.0 (82)	31.7 (82)	30.5 (82)	28.0 (82)	0.94
Non-mastitic	47.1 (136)	45.5 (68)	48.5 (68)	0.73	50.0 (34)	55.9 (34)	44.1 (34)	38.2 (34)	0.50
Escherichia coli	22.5 (40)	25 (20)	25.0 (20)	0.70	20.0 (10)	20.0 (10)	20.0 (10)	30.0 (10)	0.93
Klebsiella spp.	0 (32)	0 (16)	0 (16)	-	0 (8)	0 (8)	0 (8)	0 (8)	-
Streptococcus spp.	20.0 (120)	21.7 (60)	18.3 (60)	0.64	13.3 (30)	16.7 (30)	26.7 (30)	23.3 (30)	0.55

[1]PF: PowerFood; [2]PS: PowerSoil; [3]WM: whole milk; [4]F+P: fat + pellet

B

	Negative 16S rRNA gene amplification % (n = number of samples)										
		PowerFood					PowerSoil				
Culture groups	Total	WM[1]	Fat	F+P[2]	Pellet	*P*-value	WM	Fat	F+P	Pellet	*P*-value
Combined groups	29.6 (328)	29.3 (41)	31.7 (41)	29.3 (41)	29.3 (41)	0.99	26.8 (41)	31.7 (41)	31.7 (41)	26.8 (31)	0.92
Non-mastitic	47.1 (136)	52.9 (17)	52.9 (17)	41.2 (17)	35.3 (17)	0.84	47.1 (17)	58.2 (17)	47.1 (17)	41.2 (17)	0.77
Escherichia coli	22.5 (40)	20.0 (5)	20.0 (5)	20.0 (5)	40.0 (5)	1.0	20.0 (5)	20.0 (5)	20.0 (5)	20.0 (5)	1.0
Klebsiella spp.	0 (32)	0 (4)	0 (4)	0 (4)	0 (4)	-	0 (4)	0 (4)	0 (4)	0 (4)	-
Streptococcus spp.	20.0 (120)	13.3 (15)	20.0 (15)	26.7 (15)	26.7 (15)	0.78	13.3 (15)	13.3 (15)	26.7 (5)	20.0 (15)	0.74

WM[1]: whole milk; F+P[2]: fat + pellet

Table 2.2. Comparison of DNA and amplicon concentration among DNA extraction protocols using different milk sample types (Whole Milk, Fat, Fat + Pellet, and Pellet) according to milk-health status groups (non-mastitic or mastitis caused by *Escherichia coli*, *Klebsiella* spp., *Streptococcus* spp., and all groups combined) and two different DNA extraction kits. Numbers in parentheses represent the standard error of the mean. [a,b,c] Different superscripts between values indicate a significant difference (numbers in bold).

	DNA (ng/µl)					Amplicon (ng/µl)				
Milk-health status	WM[1]	Fat	F+P[2]	Pellet	*P*-value	WM	Fat	F+P	Pellet	*P*-value
PowerFood										
Non-mastitic	11.3 (12.8)	22.7 (12.8)	16.1 (12.8)	17.7 (12.8)	0.93	16.8 (6.3)	25.5 (6.6)	25.8 (5.7)	37.3 (5.4)	0.11
Escherichia coli	101.0 (54.5)	77.5 (54.5)	172.8 (54.5)	119.9 (54.5)	0.65	31.6 (13.4)	29.6 (13.4)	51.9 (13.4)	44.7 (15.5)	0.70
Klebsiella spp.	273.1 (52.5)	317.7 (52.5)	283.4 (52.5)	269.0 (52.5)	0.90	**63.3 (5.0)[a]**	**82.9 (5.0)[b]**	**62.3 (5.0)[a]**	**78.4 (5.0)[ab]**	**0.025**
Streptococcus spp.	201.6 (36.5)	245.1 (37.8)	235.1 (36.5)	203.4 (36.5)	0.78	41.6 (5.0)	55.2 (5.0)	47.2 (5.2)	52.8 (5.2)	0.24
Total	177.45 (19.5)	203.1 (20.9)	205.1 (20.5)	203.2 (20.2)	0.66	38.7 (3.2)	48.9 (3.4)	46.5 (3.2)	54.2 (3.3)	0.30
PowerSoil										
Non-mastitic	7.9 (7.9)	12.9 (13.1)	11.7 (13.1)	32.6 (13.5)	0.57	11.5 (5.6)	13.7 (6.0)	23.0 (5.3)	29.8 (5.1)	0.07
Escherichia coli	154.0 (61.0)	120.2 (61.0)	225.6 (61.0)	155.5 (61.0)	0.61	22.6 (11.0)	36.3 (11.0)	37.2 (11.0)	38.8 (11.0)	0.71
Klebsiella spp.	183.2 (54.0)	250.4 (54.0)	345.1 (54.0)	274.9 (62.3)	0.26	47.5 (6.7)	65.1 (6.7)	70.9 (6.7)	70.3 (7.7)	0.10
Streptococcus spp.	88.9 (35.9)	179.3 (37.1)	186.1 (38.4)	185.0 (37.1)	0.17	**25.0 (4.6)[a]**	**37.2 (4.6)[ab]**	**48.6 (5.2)[b]**	**49.5.9 (4.7)[b]**	**0.002**
Total	**99.9 (19.6)[a]**	**151.6 (19.8)[ab]**	**196.7 (20.2)[b]**	**163.7 (19.5)[ab]**	**0.04**	**27.1 (3.3)[a]**	**37.2 (3.3)[ab]**	**46.0 (3.4)[b]**	**46.4 (3.3)[b]**	**0.001**
All groups combined	**139.0 (14.4)[a]**	**176.1 (14.9)[ab]**	**201.1 (14.9)[b]**	**183.1 (14.6)[ab]**	**0.017**	**33.0 (2.4)[a]**	**42.9 (2.4)[b]**	**46.3 (2.4)[b]**	**50.4 (2.4)[b]**	**<0.0001**

WM[1]: Whole Milk; F+P[2]: Fat + Pellet

Table 2.3. Comparison of DNA and amplicon concentration between DNA extraction kits (PorwerFood and PowerSoil) according to milk sample type (whole milk, fat, fat + pellet, pellet) and milk-health status groups (non-mastitic or mastitis caused by *Escherichia coli, Klebsiella* spp., *Streptococcus* spp. and all groups combined). Numbers in parentheses represent the standard error of the mean. [a,b,c] . Differences with a value of $P \leq 0.05$ were considered significant (numbers in bold).

	DNA (ng/µl)			Amplicon (ng/µl)		
Milk health status	**PF[1]**	**PS[2]**	***P*-value**	**PF**	**PS**	***P*-value**
Whole Milk						
Non-mastitic	24.8 (24.8)	1.81 (1.81)	0.65	16.8 (5.9)	11.5 (5.9)	0.52
Escherichia coli	181.0 (52.2)	155.2 (60.3)	0.74	31.6 (8.9)	27.4 (10.2)	0.75
Klebsiella spp.	273.1 (52.2)	183.2 (52.2)	0.22	63.3 (8.8)	47.5 (10.2)	0.20
Streptococcus spp.	**232.4 (29.0)**	**109.1 (29.0)**	**0.03**	**41.6 (5.1)**	**25.0 (4.9)**	**0.01**
Total	**177.4 (19.5)**	**99.9 (19.6)**	**0.04**	**38.7 (3.2)**	**27.1 (3.3)**	**0.01**
Fat						
Non-mastitic	44.9 (42.6)	31.3 (31.3)	0.81	24.1 (6.7)	14.5 (6.7)	0.30
Escherichia coli	134.8 (52.2)	94.3 (52.2)	0.58	29.6 (8.9)	36.3 (8.9)	0.58
Klebsiella spp.	317.7 (52.2)	250.4 (52.2)	0.36	82.9 (8.8)	65.1 (8.8)	0.15
Streptococcus spp.	280.2 (31.5)	206.8 (29.0)	0.08	**54.6 (5.3)**	**37.2 (4.9)**	**0.01**
Total	203.1 (20.9)	151.6 (19.9)	0.07	**48.9 (3.4)**	**37.2 (3.3)**	**0.01**
Fat + Pellet						
Non-mastitic	31.5 (31.5)	16.6 (16.6)	0.77	25.8 (5.3)	23.0 (5.6)	0.70
Escherichia coli	269.7 (52.2)	199.9 (60.3)	0.17	51.9 (8.8)	43.7 (10.2)	0.61
Klebsiella spp.	283.4 (52.2)	345.1 (52.2)	0.40	62.3 (8.8)	70.9 (8.8)	0.49
Streptococcus spp.	257.4 (31.5)	258.1 (33.0)	0.42	47.2 (5.3)	48.6 (5.6)	0.85
Total	205.1 (20.5)	196.7 (20.2)	0.76	46.5 (3.2)	46.0 (3.4)	0.90
Pellet						
Non-mastitic	32.2 (32.2)	2.1 (2.1)	0.64	37.3 (5.1)	25.0 (5.6)	0.10
Escherichia coli	253.9 (60.3)	145.4 (52.2)	0.38	44.7 (10.2)	38.8 (8.8)	0.66
Klebsiella spp.	268.9 (52.2)	248.9 (52.2)	0.78	78.4 (8.8)	70.3 (10.2)	0.54
Streptococcus spp.	274.5 (33.0)	238.3 (31.5)	0.98	52.6 (5.6)	49.5 (5.3)	0.40
Total	203.2 (20.2)	163.7 (19.5)	0.15	54.2 (3.3)	46.4 (3.3)	0.08
All groups combined	**170.0 (9.2)**	**139 (9.2)**	**0.007**	**47.0 (1.8)**	**39.0 (1.9)**	**0.001**

PF[1], PowerFood; PS[2], PowerSoil

Post-sequencing outcomes

Overall sequencing quality results

The total post-quality-control number of sequences (sequences were filtered for size, quality, phiX reads, and for the presence of chimeras) used in the study was 15,177,739. The average coverage of sequences per sample was 68,061 (median = 57,844 sequences) with a standard deviation (SD) of 38,261.

Number of sequences and number of OTUs

Although differences between amplicon concentrations were detected among the milk fractions and extraction kits tested, sufficient amplicon yields were recovered from all samples for amplicon sequencing. The post-sequencing data after quality analysis and removal of low-quality sequences are described in Table 2.4 and Table 2.5. No significant differences between milk fractions (*P*-value = 0.77, Table 2.4) and DNA extraction kits (*P*-value = 0.48, Table 2.5) were observed in the number of 16S rRNA sequences that were generated from amplicons of non-mastitic and mastitic milk samples. Likewise, milk-health status and milk fraction did not influence the number of OTUs obtained by using either of the two DNA extraction kits (*P*-value = 0.26, Table 2.5).

Table 2.4. Comparison of the number of sequences and operational taxonomic units (OTUs) between milk sample types (whole milk, fat, fat + pellet, and pellet) according to milk-health groups (non-mastitic, mastitis caused by *Escherichia coli*, *Klebsiella* spp., *Streptococcus* spp. infection, and all groups combined) and DNA extraction kits (PorwerFood and PowerSoil). The numbers in parentheses indicate the standard error of the mean. [a,b,c] Different superscripts between values indicate a significant difference.

	Sequences (n)					OTUs (n)				
Culture groups	WM[1]	Fat	F+P[2]	Pellet	*P*-value	WM	Fat	F+P	Pellet	*P*-value
PowerFood										
Non-mastitic	92,449.8 (14,183.6)	96,982.9 (15,044.0)	62,067.4 (12,829.5)	78,636.7 (12,283.3)	0.27	55,707.9 (12,236.1)	48,676.4 (10.857.0)	32,947 (9,258.9)	45,156.2 (8,864.7)	0.69
Escherichia coli	77,462.0 (19,135.1)	85,728.2 (19135.1)	68,356.2 (19,135.1)	81,823.0 (22,095.3)	0.92	24,221.5 (3,907.0)	27,238.5 (3,907.0)	18,870.7 (3,907.0)	25,868.7 (4,511.4)	0.48
Klebsiella spp.	72,259.7 (21,875.0)	48,780.5 (21,875.0)	61657.75 (21,875.0)	72,935.5 (21,875.0)	0.84	3,1317.2 (4,932.9)	21,695.2 (4,932.9)	20,274.2 (4,932.9)	27,522.0 (4,932.9)	0.39
Streptococcus spp.	55,075.1 (9,714.1)	44,058.2 (10,560.3)	71,069.6 (10,560.3	67,644.8 (11,075.7)	0.26	27,684.1 (6,6602.2)	23,120.0 (7,177.3)	39769.0 (7,177.3)	30,626.4 (7,527.6)	0.41
Total	61,843 (7,130.2)	69,634 (7,338.7)	61,897 (7,313.9)	68,246 (7,124.6)	0.71	3,4409 (4,630.0)	26,427 (4,759.7)	29,212 (4,620.7)	31,215 (4,673.4)	0.64
PowerSoil										
Non-mastitic	69,840.6 (13,917.7)	64,387.4 (15,560.5)	72,775.2 (13,917.7)	63,337.4 (13,917.7)	0.95	41,940.2 (9,505.9)	39,374.1 (10,627.9)	26,808.4 (9,505.9)	37,011.8 (9,505.9)	0.86
Escherichia coli	81,417 (23,453.6)	89,387.5 (20,311.4)	63,141.7 (23,453.6)	79,525.7 (20,311.4)	0.86	26,271.7 (8,865.7)	30,421.2 (7.677.9)	27,526.7 (8,865.7)	31,977.2 (7,677.9)	0.95
Klebsiella spp.	67,363.5 (21,558.8)	81,043.2 (21,558.8)	60,895.5 (21,558.8)	74,181.2 (21,558.8)	0.91	24,998.2 (3,777.8)	33,842.2 (3,777.8)	24,944.7 (3,777.9)	24,489.2 (3,777.9)	0.28
Streptococcus spp.	47,555.1 (7,773.5)	56,480.1 (7,773.5)	56,752.1 (8,863.1)	71,954.6 (8,450.6)	0.22	24,126.6 (3,772.0)	22,172.5 (3,772.0)	34,084.4 (4,300.7)	27,362.5 (4,100.6)	0.19
Total	73,893 (7,618.7)	63,216 (7,832.1)	67,035 (7,603.3)	72,799 (7,690.0)	0.79	28,950 (3,853.1)	29,647 (3,965.7)	26,587 (3,952.4)	28,149 (3,850.1)	0.94
All groups combined	74,121 (7,014.3)	68,801 (7,349.6)	75,602 (7,007.0)	67,594 (7,158.7)	0.77	63,680 (7,051.2)	71,322 (7,247.7)	65,704 (7,384.2)	72,008 (7,113.0)	0.86

WM[1]: whole milk; F+P[2]: fat + pellet

Table 2.5. Comparison of the number of sequences and operational taxonomic units (OTUs) between DNA extraction kits (PowerFood and PowerSoil) according to milk-health groups (non-mastitic, mastitis caused by *Escherichia coli*, *Klebsiella* spp., *Streptococcus* spp. infection, and all groups combined) and milk sample types (whole milk, fat, fat + pellet, and pellet). Numbers in parentheses indicate the standard error of the mean. Differences with a value of $P \leq 0.05$ were considered significant.

	Sequences (n)			OTUs		
Culture groups	**PF[1]**	**PS[2]**	***P*-value**	**PF**	**PS**	***P*-value**
Whole Milk						
Non-mastitic	92,450 (12,776.0)	69,841 (12,121.0)	0.20	55,708 (7,489.1)	41,940 (7,104.7)	0.18
Escherchia coli	77,462 (19,165.0)	81,417 (2,212.09)	0.89	24,222 (11,234.0)	26,272 (12,971.0)	0.90
Klebsiella spp.	72,260 (19,165.0)	67,364 (19,165.0)	0.85	31,317 (11,234.0)	24,998 (11,234.0)	0.69
Streptococcus spp.	55,075 (10,631.0)	47,555 (10,631.0)	0.61	27,684 (6,231.3)	24127 (6,231.3)	0.68
Total	74,121 (7,014.3)	63,680 (7,051.2)	0.28	35,032 (4,135.2)	28,831 (4,157.0)	0.28
Fat						
Non-mastitic	96,983 (13,551.0)	70,540 (14,487)	0.18	48,676 (7,943.3)	43,141 (8,491.8)	0.63
Escherichia coli	85,728 (19,165.0)	89388 (19,165.0)	0.89	27,239 (11,234.0)	30,421 (11,234.0)	0.84
Klebsiella spp.	48,780 (19,165.0)	81,043 (19,165.0)	0.23	21,695 (11,234.0)	33,842 (11,234.0)	0.44
Streptococcus spp.	44,058 (11,557.0)	56,480 (10,631.0)	0.42	23,120 (6,774.1)	22,172 (6,231.3)	0.91
Total	68,801 (7,349.6)	71,322 (7,247.7)	0.80	30,113 (4,332.9)	29,850 (4,272.8)	0.96
Fat + Pellet						
Non-mastitic	62,067 (11,557.0)	72,775 (12,121.0)	0.52	1,8871 (11,234.0)	27,527 (12,971.0)	0.53
Escherichia coli	68,356 (19,165.0)	63,142 (22,129.0)	0.85	18,871 (11,234.0)	27,527 (11,234.0)	0.61
Klebsiella spp.	61,658 (19,165.0)	60,896 (19,165.0)	0.97	20,274 (11,234.0)	24,945 (11,234.0)	0.76
Streptococcus spp.	71,070 (11,557.0)	56,752 (12,121)	0.39	39,769 (6,774.1)	34,084 (7,104.7)	0.56
Total	67,594 (7,007.0)	65,704 (7,384.2)	0.71	29,952 (4,130.9)	27,322 (4,353.3)	0.65
Pellet						
Non-mastitic	78,637 (11,065.0)	63,337 (12,121)	0.35	45,156 (6,485.7)	37,012 (7,104.7)	0.39
Escherichia coli	81,823 (22,129)	79,526 (19,165)	0.93	25,869 (12,971.0)	31,977 (11,234.0)	0.72
Klebsiella spp.	72,936 (19,165.0)	74,181 (19,165.0)	0.96	27,522 (11,234.0)	24,489 (11,234.0)	0.84
Streptococcus spp.	67,645 (12,121.0)	71,955 (11,557.0)	0.79	30,626 (7,104.7)	27,363 (6,774.1)	0.73
Total	75,602 (7,158.67)	72,008 (7,113.0)	0.85	33,147 (4,220.3)	29,295 (4,193.4)	0.51
All groups combined	71,578 (3,705.5)	68,168 (3,747.6)	0.48	32,093 (2,181.0)	28,831 (2,205.8)	0.26

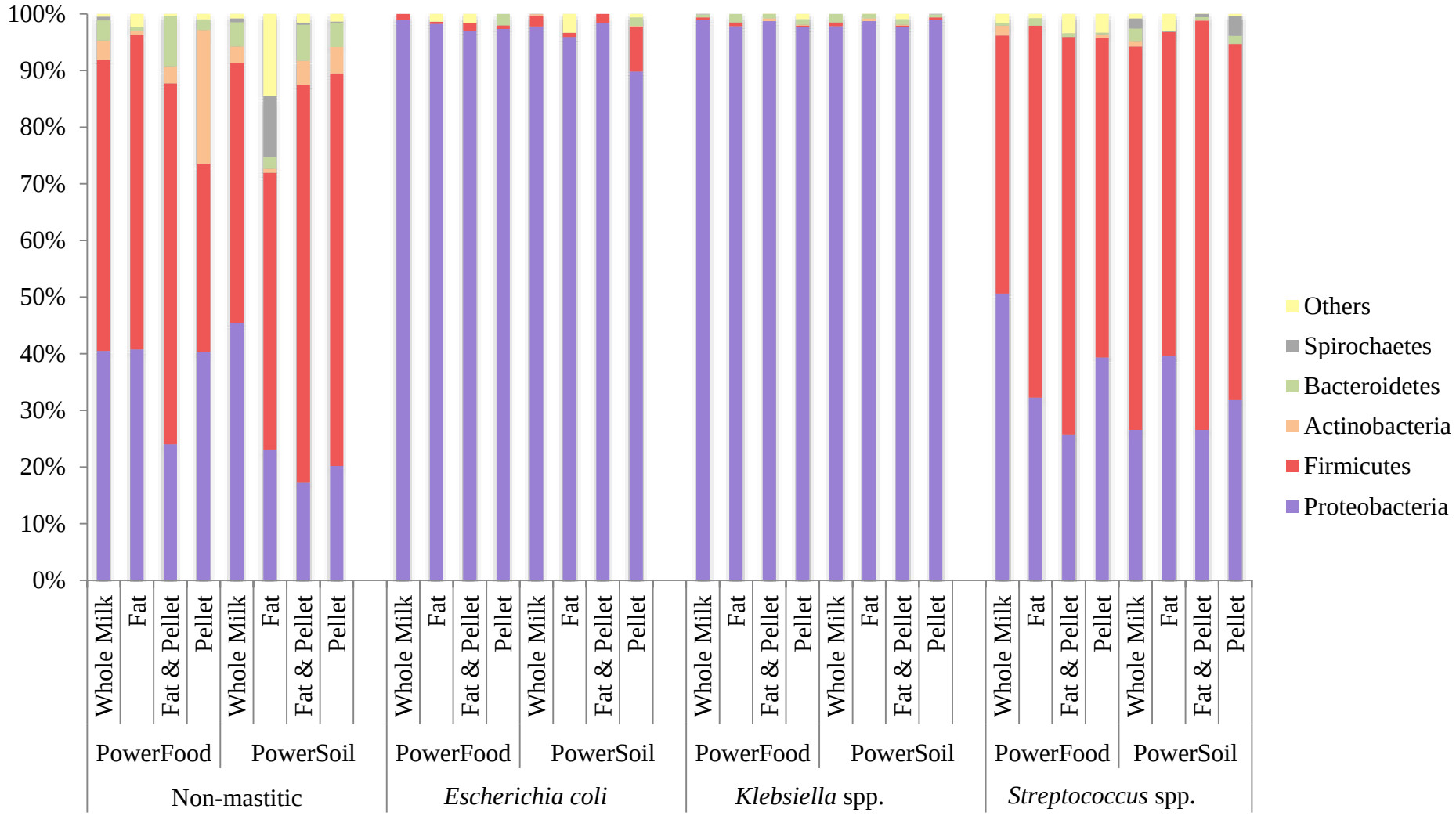
PF[1], PowerFood; PS[2], PowerSoil

Taxonomic profile at the phylum level

The on-farm culture system for bacterial identification were used as a criterion to assess the effectiveness of the DNA extraction methods on isolation of the DNA from milk microbial communities providing a foundation for the identification of the more appropriate method for milk bacterial profiling and identification of the causative agent of clinical mastitis. Here, we evaluated how the taxa detection frequency was affected by different DNA isolation procedures.

The relative distribution of the most common phyla detected in non-mastitic and mastitic samples are presented in Figure 2.2. Sequences affiliated with Firmicutes (MRA: 57.7%, standard error - SE: 7.6) and Proteobacteria (MRA: 26.0, SE: 7.6) dominated the non-mastitic milk samples, regardless of the DNA isolation method (Figure 2.2). However, in mastitic samples due to *Escherichia coli* and *Klebsiella* spp., the phylum Proteobacteria accounted for approximately 98% of the detected 16S rRNA sequences (Figure 2.2). In regards to the *Streptococcus* spp. group, the majority of the sequences were affiliated with Proteobacteria (MRA: 69.6, SE: 9.5) and Firmicutes ((MRA: 30.1, SE: 9.4, Figure 2.2), respectively. Overall, the microbial community depicted by 16S rRNA gene sequencing was dictated by milk bacterial species and not the DNA isolation method used (Figure 2.2).

Figure 2.2: Mean relative abundance of the most prevalent bacterial phyla identified in non-mastitic milk samples and milk samples from cows diagnosed with clinical mastitis due to *Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp. infection according to four milk sample types (whole milk, fat, fat + pellet, and pellet) and two different DNA extraction kits (PowerFood and PowerSoil).



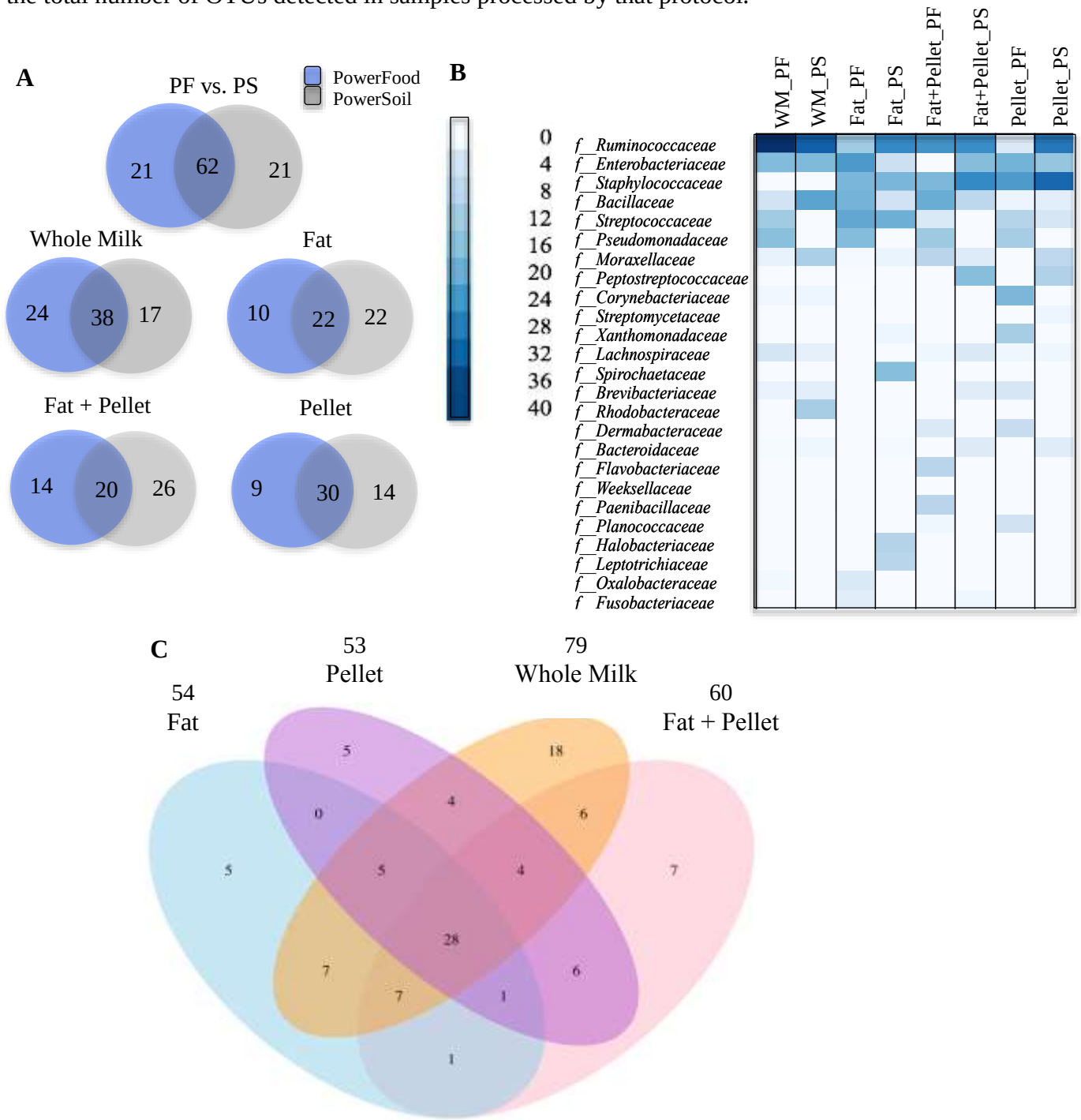
Taxonomic profile at the family level

Non-mastitic milk

Figure 2.3A displays a Venn diagram illustrating the degree of overlap of bacterial OTUs at the family level between the two DNA extraction kits for non-mastitic milk samples. The Venn diagram shows that, overall, 62 families were shared between the samples extracted by the two DNA extraction kits (Figure 2.3A). However, when the milk sample types are considered separately, the number of shared bacterial families decreased to 38, 30, 22 and 20 for whole milk, fat, fat + pellet, and pellet, respectively (Figure 2.3A). The most common bacteria families detected in non-mastitic milk samples were represented by f__Ruminococaceae, f__Enterobacteriaceae, f__Staphylococcaceae, f__Bacillaceae, f__Streptococcaceae, and f__Pseudomonadaceae (Figure 2.3B).

The degree of overlap of bacterial OTUs at the family level among the four milk sample types is described in Figure 2.3C. Twenty-eight shared OTUs comprise the core microbiota identified among all four milk sample types. A detailed description of the 28 shared families is given in Supplemental 1. f__Ruminococaceae, f__Enterobacteriaceae, f__Bacillaceae, and f__Pseudomonadaceae were the top four bacterial families identified in milk samples regardless of the milk sample type (Supplemental 2.1). A more detailed description of the types of unique families detected in each milk sample type is given in Supplemental 2.2.

Figure 2.3: Venn diagrams showing the numbers of unique and shared bacterial OTUs at the family level among non-mastitic milk samples extracted by PowerFood (PF) and PowerSoil (PS) (A). Heatmap illustrating the mean relative abundance of the 25 most common bacterial families identified in non-mastitic milk samples according to each milk fraction and DNA extraction kit. Small relative abundance values are white, progressing to higher values as dark blue. Each square in the heatmap represents the abundance level of a single category (B). Venn diagram showing the numbers of unique and shared bacterial OTUs according to milk sample type: fat, fat + pellet, pellet, and whole milk (WM) (C). The numbers at the top of the fraction name are the total number of OTUs detected in samples processed by that protocol.



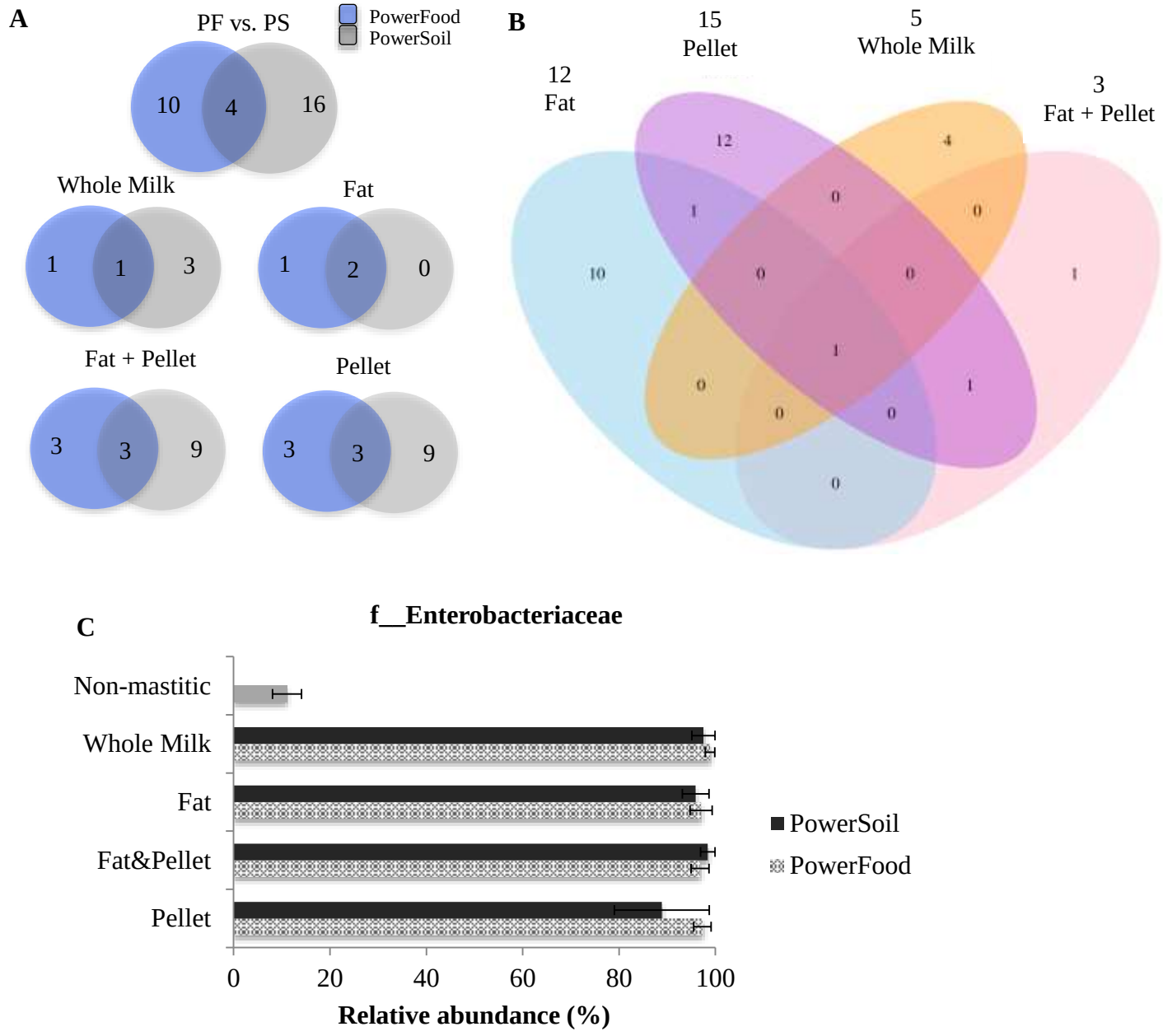
Escherichia coli mastitic milk

Figure 2.4A shows a Venn diagram illustrating the degree of overlap of bacterial OTUs at the family level between the two DNA extraction kits for mastitic milk due to *Escherichia coli*. Of the 30 OTUs detected, only four, f__Bacteroidaceae, f__Ruminococcaceae, f__Staphylococcaceae, and f__Enterobacteriaceae, were shared between the samples extracted by PowerFood and PowerSoil (Figure 2.4A).

In differentiating the milk bacterial OTUs, we also investigated the numbers of shared and unique bacteria detected in samples extracted from each milk sample type. As displayed in Figure 4B, only one family, f__Enterobacteriaceae, comprised the core microbiota in the samples from whole milk, fat, fat + pellet, and pellet. In total, 10, 12, 4 and 1 bacterial families were exclusively found in samples extracted from fat, pellet, whole milk, and fat + pellet, respectively (Figure 2.4B). Details of the unique OTUs at the family level detected in each milk sample types are shown in Supplemental 2.3.

We also evaluated whether the relative abundance of the mastitis causative agent identified by bacterial culture varied with the DNA isolation method. Figure 2.4C shows that the MRA for all OTUs affiliated to f__Enterobacteriaceae showed good reproducibility (MRA > 90%) across all milk sample type and both DNA extraction kits (Figure 2.4C). The eight DNA isolation procedures accurately detected f__Enterobacteriaceae, and no significant differences in the relative abundance of this taxon were observed among the DNA isolation procedures (P – value > 0.05, Figure 2.4C).

Figure 2.4. Venn diagram showing the numbers of unique and shared bacterial OTUs at the family level among *Escherichia coli* mastitic milk samples processed by PowerFood and PowerSoil DNA extraction kits (A), and by milk sample type (whole milk, fat, fat + pellet, and pellet) (B). The numbers at the top of the milk sample type name indicate the total number of OTUs detected in samples processed by that protocol. Mean relative abundance of the f__Enterobacteriaceae taxon among milk sample types and DNA extraction kits (C). Error bars represent the standard error of the mean and non-mastic milk group was added as a control sample.



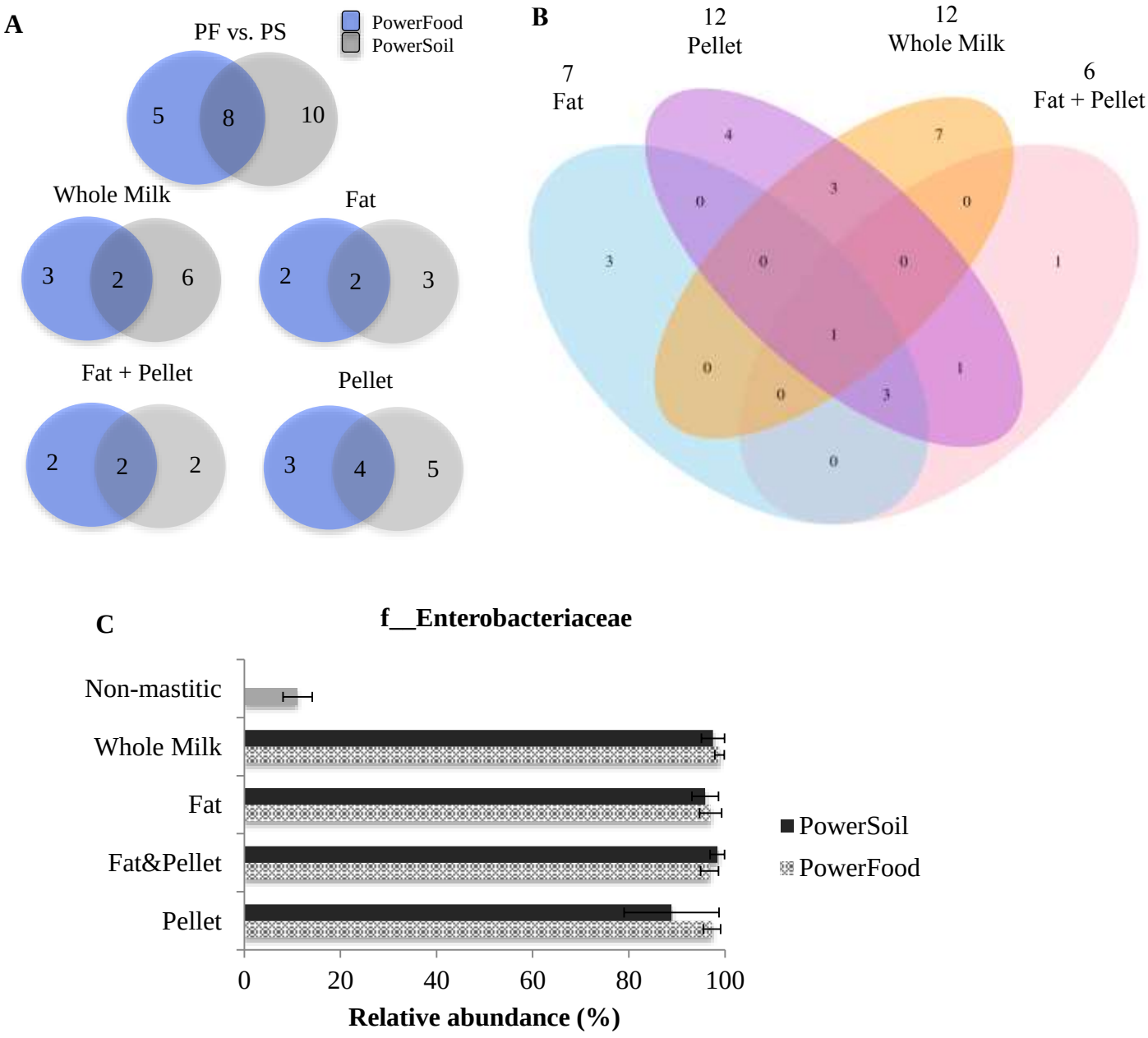
Klebsiella spp. mastitic milk

Figure 2.5A shows a Venn diagram illustrating the degree of overlap of bacterial OTUs at the family level between the two DNA extraction kits for mastitic milk samples due to *Klebsiella* spp. Of the 28 families detected, only eight, f__Halobacteriaceae, f__Corynebacteriaceae, f__Ruminococcaceae, f__Brevibacteriaceae, f__Bacteroidaceae, f__Flavobacteriaceae, f__Pseudomonadaceae, and f__Enterobacteriaceae, were shared between the samples processed by PowerFood and PowerSoil (Figure 2.5A).

On the other hand, as shown in the Venn diagram depicted in Figure 2.5B, only one family, f__Enterobacteriaceae, was shared among the samples processed from whole milk, fat, fat + pellet, and pellet. In total, 3, 4, 7, and 1 bacterial families were found to be unique to the samples extracted from fat, pellet, whole milk, and fat + pellet, respectively (Figure 2.5B). A more detailed description of the types of unique OTUs at the family level detected in each protocol is given in Supplemental 2.4.

Figure 5C shows that the MRA for all OTUs affiliated to f__Enterobacteriaceae showed good reproducibility (MRA > 90%) across all milk sample types and DNA extraction kits. The eight DNA isolation procedures accurately detected f__Enterobacteriaceae, and no significant differences in the relative abundance of this taxon were observed among the DNA isolation procedures (P - value > 0.05, Figure 2.5C).

Figure 2.5: Venn diagram showing the numbers of unique and shared bacterial OTUs at the family level among *Klebsiella* spp. mastitic milk samples processed by PowerFood and PowerSoil DNA extraction kits (A), and by milk sample type (whole milk, fat, fat + pellet, and pellet) (B). The numbers at the top of the milk sample type name indicate the total number of OTUs detected in samples processed by that protocol. Mean relative abundance of the f__Enterobacteriaceae taxon among milk sample types and DNA extraction kits (C). Error bars represent the standard error of the mean and non-mastic milk group was added as a control sample.



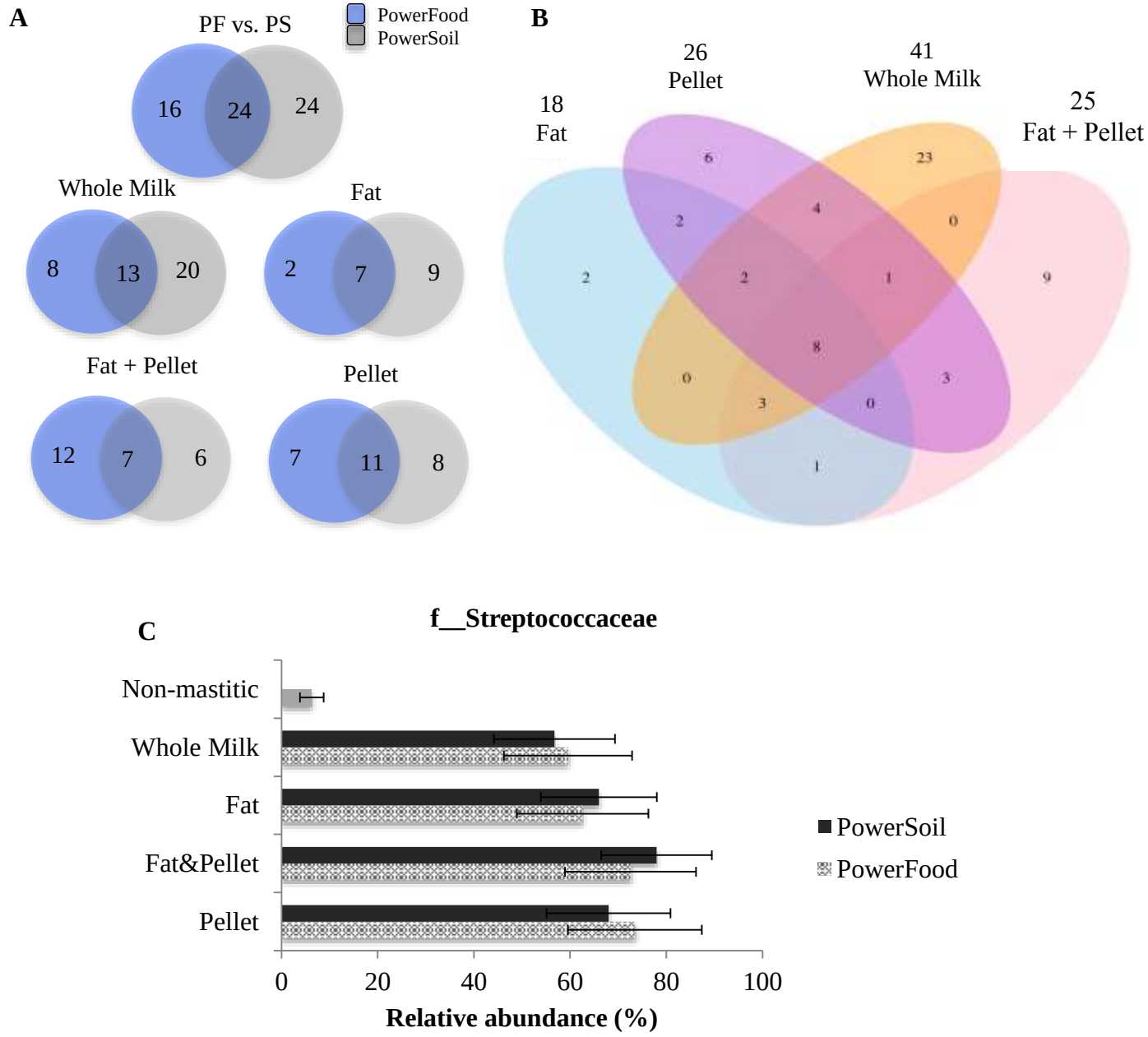
Streptococcus spp. mastitic milk

Figure 2.6 shows a Venn diagram of the degree of overlap of bacterial OTUs at the family level between the two DNA extraction kits for milk samples from cows with clinical mastitis caused by *Streptococcus* spp. infection. Of the 64 OTUs detected, 24 were shared between samples isolated by PowerFood and PowerSoil (Figure 2.6A). This number decreased to 13, 7, 7, and 7 bacterial families when considered by sample type: whole milk, pellet only, fat, and fat + pellet, respectively (Figure 2.6A).

The degree of overlap of bacterial OTUs at the family level among the four milk sample types is described in Figure 2.6B. Eight families comprise the core microbiome shared by all four milk sample types: f__Streptococcaceae, f__Pseudomonadaceae, f__Ruminococaceae, f__Enterococcaceae, f__Flavobacteriaceae, f__Enterobacteriaceae, f__Brachyspiraceae, and f__Desulfurococcaceae (Supplemental 2.5). A more detailed description of the types of unique families detected in each protocol is given in Supplemental 2.6.

Figure 2.6C shows that the MRA for all OTUs affiliated to f__Streptococcaceae had good reproducibility (MRA > 50%) across all DNA isolation methods (Figure 2.6C). The eight DNA isolation methods accurately detected f__Streptococcaceae, and no significant differences in the relative abundance of this taxon were observed between DNA isolation procedures (*P* - value > 0.05, Figure 2.6C).

Figure 2.6: Venn diagram showing the numbers of unique and shared bacterial OTUs at the family level among *Streptococcus* spp. mastitic milk samples processed by PowerFood and PowerSoil DNA extraction kits (A), and according to milk sample type (whole milk, fat, fat + pellet, and pellet) (B). The numbers at the top of the milk sample type name indicate the total number of OTUs detected in the samples processed by that protocol. Mean relative abundance of f__Streptococcaceae taxon among milk sample types and DNA extraction kits (C). Error bars represent the standard error of the mean and non-mastic milk group was added as a control sample.



Effect of DNA extraction procedures on f_ Staphylococcaceae abundance

Since, *Staphylococcus* spp. have been described as a bacterium with ability of binding to the milk fat globules (Ali-Vehmas et al., 1997), a detailed evaluation of the effect of DNA extraction methods on the relative abundance of *f__Staphylococcaceae* in milk samples from clinical and non-clinical cases of mastitis was performed.

According to our sequencing results, the family *f__Staphylococcaceae* was completely absent in the *Klebsiella* spp. and *Streptococcus* spp. mastitic groups (Supplemental Table 2.7). This scenario was not the same in *Escherichia coli* group when DNA was extracted from fat & pellet and pellet only; the MRA of this bacterium was 1.7% and 0.7% respectively, in those samples extracted by PowerFood and, 1.6% and 0.85% respectively, in samples extracted by PowerSoil (Supplemental Table 2.7). Similar low abundances were found in milk samples obtained from the non-mastitic group only when DNA extraction protocol was applied to whole milk (MRA = 0.9%) regardless of the DNA extraction kit. Despite the fact that no statistically significant differences in *f__Staphylococcaceae* relative abundance were detected for DNA extraction protocols and DNA extraction kits, the highest relative abundance of *f__Staphylococcaceae* was detected in non-mastitic milk samples submitted to DNA extraction of the pellet using PowerSoil kit (MRA = 28.6%, *P*-value > 0.05, Supplemental Table 2.7).

DISCUSSION

In the present study, we evaluated eight DNA isolation methods to compare their relative efficiency of DNA extraction from milk samples from healthy cows and cows diagnosed with clinical mastitis as well as their ability to accurately diagnose clinical mastitis due to *Escherichia coli, Klebsiella* spp. and *Streptococcus* spp. infection. Regardless of the milk sample type and the

extraction kit used, the resulting DNA and amplicon concentrations obtained from both non-mastitic and mastitic samples were adequate for next-generation sequencing. Overall, there was a fairly high level of agreement (60%) between the two DNA extraction kits (PowerFood and PowerSoil) in terms of the number of bacterial families detected in all four sample types (whole milk, fat, fat + pellet, pellet) from healthy cows; these samples shared 28 OTUs, thus defining the core microbiome framework at the family level. Furthermore, all eight DNA isolation procedures efficiently diagnosed the causative agent of mastitis (*Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp.) and uniformly characterized the most abundant bacterial families in mastic and non-mastitic milk samples.

Our data demonstrate that all eight DNA isolation methods gave adequate amounts of DNA from mastitic samples, regardless of the mastitis causative agent (*Escherichia coli, Klebsiella* spp. and *Streptococcus* spp. groups); however, the non-mastitic samples yielded relatively low amounts of DNA. These results were confirmed by gel electrophoresis, where 47.1% of the non-mastitic samples failed to produce sharp DNA bands. Additionally, the number of samples that failed (based on gel electrophoresis) did not differ among any of the eight DNA isolation protocols. These results were expected for non-mastitic milk, which typically has a low bacterial load (Zadoks et al., 2014).

Generally, the milk samples extracted by PowerFood had higher DNA concentrations than those extracted by PowerSoil, particularly the samples extracted from whole milk. Assessment of the 16S rRNA amplicon concentration revealed a similar result, in which samples extracted by PowerFood had higher amplicon concentrations than those extracted by PowerSoil These findings might be ascribed to differences in bead size between the two DNA extraction kits: the PowerFood kit contains a 0.15-mm garnet beads whereas the PowerSoil kit contains

round garnet beads approximately 0.7 mm in diameter (MoBio, Tips & Tricks ebook). For instance, in a recent study aimed at optimizing the recovery of bacterial DNA from coral habitat, Weber and colleagues (2017) evaluated four DNA extraction kits from MoBio containing distinct bead sizes and found differences in DNA concentration values, which they attributed to differences in bead size and the duration of mechanical lysis (Weber et al., 2017). It has been reported that larger beads are more likely to lyse larger cells (Shinzato et al., 2014), whereas smaller beads are more likely to target smaller microbial cells (Koonin and Wolf, 2008).

Additionally, in our study, the duration of the mechanical cell lysis procedure was similar for all samples: however, following the manufacturer's instructions, the samples extracted by the PowerSoil kit underwent a heating step at 70 ºC for optimization of cell lysis, followed by extra cleanup steps. In contrast, when using the PowerFood kit, the bacterial cells were lysed at 65 ºC. Such heating causes protein denaturation and aggregation (Dubois et al., 1991). Moreover, nearly any organic molecule will degrade at very high temperatures, and some will degrade at lower temperatures due to unstable structures or heat-labile bonds (Vieille and Zeikus, 2001). Although both PowerFood and PowerSoil were designed for isolation of microbial DNA, the larger size of PowerSoil beads in addition to the higher temperature in the cell lysis step might have affected its ability to recover bacterial DNA from milk compared to the PowerFood kit.

To ensure that the DNA extraction methods evaluated in this study would properly and accurately describe the bacterial community present in non-mastitic and mastitic milk samples, we subjected normalized amplicon libraries to DNA sequencing. No significant differences were observed in the number of sequences and numbers of OTUs among the eight DNA methods, regardless of whether the comparison was made within samples from clinical mastitis cases (i.e., due to *Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp. infection) or within samples from

healthy cows.

Milk samples from healthy cows were dominated by the Firmicutes and Proteobacteria phyla, and the most abundant bacterial families were represented by f__Ruminococcaceae and f__Enterobacteriaceae. These results are in agreement with a previous study from our group, in which the microbial profile of milk samples from healthy quarters was determined (Ganda et al., 2016b). Additionally, when we investigated potential extraction bias for the DNA isolation methods that we tested, microbial community analysis showed that 16S rRNA sequences detected within the milk samples of the *Escherichia coli* and *Klebsiella* spp. mastitis groups were dominated by Proteobacteria, whereas the samples from the *Streptococcus* spp. mastitis group were dominated by Firmicutes and Proteobacteria.

Taxonomic data at the family level showed that sequences from mastitic milk samples cultured positive for *Escherichia coli* and *Klebsiella* spp. were predominantly affiliated with f__Enterobacteriaceae regardless of the DNA isolation method used. By contrast, sequences from mastitic milk cultured positive for *Streptococcus* spp. were dominated by f__Streptococcacea, but with a goodly proportion of f__Pseudomonadaceae and f__Enterococcaceae, belonging to the phylum Proteobacteria. The relative abundance f__Streptococcacea within samples from the *Streptococcus* spp. mastitic group showed a wider distribution (50% to 80%), but the values are not significantly different between DNA isolation protocols when compared to the higher detection of f__Enterobacteriaceae in the mastitic samples resulting from infection by *Escherichia coli* and *Klebsiella* spp. These results might be explained by the distinct characteristics of mammary gland infection caused by a coliform, gram-negative bacterium (e.g. *Escherichia coli* and *Klebsiella* spp.) versus a gram-positive bacterium (e.g. *Streptococcus* spp.) (Oviedo-Boyso et al., 2007, Wellnitz and Bruckmaier, 2012). Coliform

mastitis typically leads to a more severe and aggressive infection accompanied by a high bacterial growth rate in the mammary gland compared to infections caused by some gram-positive bacteria (Wellnitz and Bruckmaier, 2012), which might be the case for *Streptococcus* spp. mastitis infections.

Although microbial community analysis revealed that most of the microbial community composition corresponded to milk bacterial species irrespective of the DNA isolation method, it is noteworthy that milk-health status group and DNA extraction kit affected the detection of f__Staphylococcaceae. For the milk-health status groups non-mastitic and *Escherichia coli* mastitis, in which f__Staphylococcaceae was present, the protocols that used whole milk were not able to accurately extract DNA from bacteria of this taxon. Furthermore, the protocol that used fat was not the best protocol for its detection. In a previous study by Ali-Vehmas et al (1997), the authors provided evidence of *Staphylococcus aureus* binding to milk fat globules (Ali-Vehmas et al., 1997). However, in our study, the highest relative abundance of f__Staphylococcaceae was detected in non-mastitic milk samples subjected to DNA extraction from pellet and not from fat; thus, the f__Staphylococcaceae sequences identified in our study are more likely to be affiliated to other *Staphylococcus* species besides *Staphylococcus aureus*.

When we investigated the level of agreement between the two DNA extraction kits in their accuracy for extracting bacterial DNA from a variety of bacterial cell morphologies, we found that 62 families comprise the common microbial core present in non-mastitic milk samples, indicating a level of agreement of 60%. By contrast, only 13%, 35% and 37% of the OTUs detected in the *Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp. mastitis groups, respectively, were shared between the microbial communities detected by PowerFood and PowerSoil. The high disagreement levels detected in this study between the DNA extraction kits

are a reflection of the kits' inability in detect rare OTUs (low abundant OTUs at the family level) and not their ability to detect the dominant bacteria, because the causative pathogen of clinical mastitis was accurately detected and showed to dominate the environment (the MRA of the mastitis pathogens was greater than 50%), regardless of the mastitis group evaluated. Data reported by Weber and colleagues also showed the ability of certain MoBio kits to accurately describe the most common bacterial species in coral samples, and their disagreement in the detection of rare OTUs (Weber et al., 2017).

Concerning the level of agreement between bacterial profiles among the DNA isolation procedures, only 27% (28/104) of the OTUs detected in non-mastitic milk were shared among the samples extracted from all four types of milk samples (whole milk, fat, fat + pellet and pellet). The percentages dropped to 3% (1/30), 4% (1/23), and 12% (8/64) for the samples from cows with mastitis caused by *Escherichia coli*, *Klebsiella* spp., and *Streptococcus* spp. infection, respectively. Similar to the results for the DNA extraction kits, all protocols appropriately detected the dominant bacterial families, regardless of milk-health status; however, the detection of rare OTUs was evidently affected by the type of protocol applied to the milk sample. Thus, if the goal of the investigation is to detect specific or rare bacterial microorganisms, care should be taken when choosing the milk sample type and DNA extraction kit for analysis of clinical and non-clinical cases of mastitis.

Given the relative successes of all eight DNA isolation methods in uniformly generating the post-sequencing outcomes for both non-mastitic and mastitic samples, the most convenient approach for extracting metagenomic DNA from milk samples would be the PowerFood kit applied to whole milk. A significant advantage of using whole milk is that it avoids the need of higher volumes of milk and the additional time and labor required for milk fractionation.

Additionally, the PowerFood kit is more cost effective and requires fewer processing steps than the PowerSoil kit. On the whole, our results suggest that whichever of the eight DNA isolation methods is chosen, it will not significantly bias the composition of the bacterial community if the aim of the study is to elucidate large differences in bacterial composition that correspond to changes in mammary gland health. Since studies that investigate the impact of mammary gland infection on milk bacterial profile usually focus on larger comparisons and differences, we recommend the use of PowerFood and whole milk for broad investigations of milk microbial dynamics.

CONCLUSION

The metagenomic DNA extracted from whole milk, fat, fat + pellet or pellet by using either a PowerFood or PowerSoil kit was of sufficient quality and concentration to enable detection of the causative agents of mastitis (*Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp.). All eight DNA isolation methods tested here recovered DNA suitable for PCR amplification of the 16S rRNA gene from mastitic milk samples. The DNA isolation procedures still need to be improved for more efficient isolation of metagenomic DNA from non-mastitic milk. Differences in the numbers and types of unique OTUs were observed among the DNA isolation methods regardless of milk-health status group (non-mastitic and mastitic); however these were rare OTUs (i.e., low in abundance). Taxonomic analysis revealed that, on a broad scale, overall microbial community structure was significantly determined by the milk bacterial community rather than by the milk sample type or DNA extraction kit, a result that validates the use of either the PowerFood or PowerSoil kit as well as whole milk, fat, fat + pellet or pellet for broad comparisons of the milk microbiota, regardless of whether the milk sample is from a clinical case of mastitis or from healthy cows. Based upon these results and the labor/cost-

benefits, we suggest the use of PowerFood and whole milk for DNA extraction for cultivation-independent characterization of non-mastitic and mastitic milk bacterial communities.

ACKNOWLEDGEMENTS

This project was supported by the Agriculture and Food Research Initiative Competitive Grant no. 2013-67015-21233 from the USDA National Institute of Food and Agriculture (Washington, DC.).

REFERENCES

Ali-Vehmas, T., P. Westphalen, V. Myllys, and M. Sandholm. 1997. Binding of Staphylococcus aureus to milk fat globules increases resistance to penicillin-G. J Dairy Res 64(2):253-260.

Bag, S., B. Saha, O. Mehta, D. Anbumani, N. Kumar, M. Dayal, A. Pant, P. Kumar, S. Saxena, K. H. Allin, T. Hansen, M. Arumugam, H. Vestergaard, O. Pedersen, V. Pereira, P. Abraham, R. Tripathi, N. Wadhwa, S. Bhatnagar, V. G. Prakash, V. Radha, R. M. Anjana, V. Mohan, K. Takeda, T. Kurakawa, G. B. Nair, and B. Das. 2016. An improved method for high quality metagenomics DNA extraction from human and environmental samples. Sci Rep 6:26775.

Batt C.A. and Tortorello ML. 2014. Encyclopedia of food microbiology. Vol. 1. Second ed. Microbiology of cream and butter. Elsevier.

Beaver, A., C. L. Cazer, P. L. Ruegg, Y. T. Gröhn, and Y. H. Schukken. 2016. Implications of PCR and ELISA results on the routes of bulk-tank contamination with Mycobacterium avium ssp. paratuberculosis. J Dairy Sci 99(2):1391-1405.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13(7):581-583.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., Knight, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8):1621-1624.

Dalgleish, D. G. and M. Corredig. 2012. The structure of the casein micelle of milk and its changes during processing. Annu Rev Food Sci Technol 3:449-467.

Dubois, M. F., A. G. Hovanessian, and O. Bensaude. 1991. Heat-shock-induced denaturation of proteins. Characterization of the insolubilization of the interferon-induced p68 kinase. J Biol Chem 266(15):9707-9711.

Federation, I. D. 1987. Machine Milking and Mastitis. Bulletin of the International Dairy Federation

Ganda, E. K., R. S. Bisinotto, D. H. Decter, and R. C. Bicalho. 2016a. Evaluation of an on-farm culture system (Accumast) for fast identification of milk pathogens associated with clinical mastitis in dairy cows. PLoS One 11(5):e0155314.

Ganda, E. K., R. S. Bisinotto, S. F. Lima, K. Kronauer, D. H. Decter, G. Oikonomou, Y. H. Schukken, and R. C. Bicalho. 2016b. Longitudinal metagenomic profiling of bovine milk to assess the impact of intramammary treatment using a third-generation cephalosporin. Sci Rep 6:37565.

Gilbert, J. A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C. T., Desai, N., Eisen, J. A., Evers, D., Field, D., W. Feng, Huson, D., Jansson, J., Knight, R., Knight, J., Kolker, E., Konstantindis, K., Kostka, J., Kyrpides, N., and R. Mackelprang, McHardy, A., Quince, C., Raes, J., Sczyrba, A., Shade, A., Stevens, R. 2010. Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. Stand Genomic Sci 3(3):243-248.

Hagnestam-Nielsen, C. and S. Ostergaard. 2009. Economic impact of clinical mastitis in a dairy herd assessed by stochastic simulation using different methods to model yield losses. Animal 3(2):315-328.

Herthnek, D., S. S. Nielsen, A. Lindberg, and G. Bölske. 2008. A robust method for bacterial lysis and DNA purification to be used with real-time PCR for detection of Mycobacterium avium subsp. paratuberculosis in milk. J Microbiol Methods 75(2):335-340.

Hickey, C. D., J. J. Sheehan, M. G. Wilkinson, and M. A. Auty. 2015. Growth and location of bacterial colonies within dairy foods using microscopy techniques: a review. Front Microbiol 6:99.

Koonin, E. V. and Y. I. Wolf. 2008. Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. Nucleic Acids Res 36(21):6688-6719.

Koskinen, M. T., J. Holopainen, S. Pyörälä, P. Bredbacka, A. Pitkälä, H. W. Barkema, R. Bexiga, J. Roberson, L. Sølverød, R. Piccinini, D. Kelton, H. Lehmusto, S. Niskala, and L. Salmikivi. 2009. Analytical specificity and sensitivity of a real-time polymerase chain reaction assay for identification of bovine mastitis pathogens. J Dairy Sci 92(3):952-959.

Lago, A., S. M. Godden, R. Bey, P. L. Ruegg, and K. Leslie. 2011. The selective treatment of clinical mastitis based on on-farm culture results: II. Effects on lactation performance, including clinical mastitis recurrence, somatic cell count, milk production, and cow survival. J Dairy Sci 94(9):4457-4467.

Levison, L. J., E. K. Miller-Cushon, A. L. Tucker, R. Bergeron, K. E. Leslie, H. W. Barkema, and T. J. DeVries. 2016. Incidence rate of pathogen-specific clinical mastitis on conventional and organic Canadian dairy farms. J Dairy Sci 99(2):1341-1350.

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., Andersen, G. L., Knight, R., Hugenholtz, P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610-618.

National Mastitis Council. 1998. Current Concepts of Bovine Mastitis. 4 ed. Madison, WI, Natl. Mastitis Council

NM, C. 1999. Laboratory Handbook on Bovine Mastitis: National Mastitis Council. Page 222.

Oliveira, L., C. Hulland, and P. L. Ruegg. 2013. Characterization of clinical mastitis occurring in cows on 50 large dairy herds in Wisconsin. J Dairy Sci 96(12):7538-7549.

Oultram, J. W., E. K. Ganda, S. C. Boulding, R. C. Bicalho, and G. Oikonomou. 2017. A Metataxonomic Approach Could Be Considered for Cattle Clinical Mastitis Diagnostics. Front Vet Sci 4:36.

Oviedo-Boyso, J., J. J. Valdez-Alarcón, M. Cajero-Juárez, A. Ochoa-Zarzosa, J. E. López-Meza, A. Bravo-Patiño, and V. M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J Infect 54(4):399-409.

Phuektes, P., G. F. Browning, G. Anderson, and P. D. Mansell. 2003. Multiplex polymerase chain reaction as a mastitis screening test for Staphylococcus aureus, Streptococcus

agalactiae, Streptococcus dysgalactiae and Streptococcus uberis in bulk milk samples. J Dairy Res 70(2):149-155.

Pyörälä, S. 2009. Treatment of mastitis during lactation. Ir Vet J 62 Suppl 4:S40-44.

Quigley, L., R. McCarthy, O. O'Sullivan, T. P. Beresford, G. F. Fitzgerald, R. P. Ross, C. Stanton, and P. D. Cotter. 2013. The microbial content of raw and pasteurized cow milk as determined by molecular approaches. Journal of Dairy Science 96(8):4928-4937.

Quigley, L., O. O'Sullivan, T. P. Beresford, R. Paul Ross, G. F. Fitzgerald, and P. D. Cotter. 2012. A comparison of methods used to extract bacterial DNA from raw milk and raw milk cheese. J Appl Microbiol 113(1):96-105.

Rasolofo, E. A., D. St-Gelais, G. LaPointe, and D. Roy. 2010. Molecular analysis of bacterial population structure and dynamics during cold storage of untreated and treated milk. Int J Food Microbiol 138(1-2):108-118.

Shinzato, C., S. Mungpakdee, N. Satoh, and E. Shoguchi. 2014. A genomic approach to coral-dinoflagellate symbiosis: studies of Acropora digitifera and Symbiodinium minutum. Front Microbiol 5:336.

Steiner, J. J., C. J. Poklemba, R. G. Fjellstrom, and L. F. Elliott. 1995. A rapid one-tube genomic DNA extraction process for PCR and RAPD analyses. Nucleic Acids Res 23(13):2569-2570.

Taponen, S., L. Salmikivi, H. Simojoki, M. T. Koskinen, and S. Pyörälä. 2009. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. J Dairy Sci 92(6):2610-2617.

USDA. 2014. Dairy 2014: Milk Quality, Milking Procedures, and Mastitis on U.S. Dairies. USDA, Animal and Plant Health Inspection Service, Veterinary Services, Center for Epidemiology and Animal Health, Fort Collins, CO.

Vieille, C. and G. J. Zeikus. 2001. Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65(1):1-43.

Walstra, P., Geurts, T.J., Noomen, A., Jellema, A., and van Boekel, M.A.J.S. 1999. Dairy Technology: Principles of Milk Properties and Processes. Marcel Dekker, Inc., New York.

Weber, L., E. DeForce, and A. Apprill. 2017. Optimization of DNA extraction for advancing coral microbiota investigations. Microbiome 5(1):18.

Wellnitz, O. and R. M. Bruckmaier. 2012. The innate immune response of the bovine mammary gland to bacterial infection. Vet J 192(2):148-152.

Wilson, I. G. 1997. Inhibition and facilitation of nucleic acid amplification. Appl Environ Microbiol 63(10):3741-3751.

Zadoks, R. N., R. Tassi, E. Martin, J. Holopainen, S. McCallum, J. Gibbons, and K. T. Ballingall. 2014. Comparison of bacteriological culture and PCR for detection of bacteria in ovine milk--sheep are not small cows. J Dairy Sci 97(10):6326-6333.

Supplemental Table 2.1: Description of the core microbiota, characterized as OTUs at the family level detected in all milk sample types (whole milk, fat, fat + pellet, and pellet) of non-mastitic milk samples.

	Mean relative abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Pseudomonadaceae	5.7 (5.7)	6.5 (6.4)	5.3 (5.3)	5.0 (5.0)
f__Moraxellaceae	5.3 (3.8)	1.3 (0.4)	4.6 (2.7)	2.7 (2.4)
f__Rhodobacteraceae	4.9 (4.7)	0.02 (0.02)	0.2 (0.1)	0.2 (0.1)
f__Streptococcaceae	4.9 (4.6)	16.2 (9.0)	1.4 (1.2)	5.3 (4.0)
f__Ruminococcaceae	32.9 (7.9)	16.6 (7.1)	22.4 (7.3)	13.2 (5.8)
f__Lachnospiraceae	2.5 (0.8)	0.7 (0.5)	1.9 (0.7)	1.4 (0.05)
f__Enterobacteriaceae	13.1 (7.1)	12.4 (6.4)	6.4 (4.2)	12.9 (6.6)
f__Bacillaceae	11.2 (4.7)	9.2 (6.0)	11.3 (5.5)	1.7 (0.7)
f__Bacteroidaceae	1.3 (0.4)	0.7 (0.7)	1.5 (0.6)	1.3 (0.4)
f__Flavobacteriaceae	0.09 (0.08)	0.3 (0.3)	3.6 (0.6)	0.2 (0.1)
f__Clostridiaceae	0.8 (0.3)	0.8 (0.4)	0.2 (0.1)	0.1 (0.07)
f__Oxalobacteraceae	0.7 (0.4)	1.6 (1.0)	0.5 (0.5)	0.2 (0.2)
f__Corynebacteriaceae	1.4 (0.9)	0.2 (0.2)	0.2 (0.2)	7.9 (4.8)
f__Paraprevotellaceae	0.6 (0.4)	0.01 (0.01)	0.3 (0.1)	0.1 (0.1)
f__Spirochaetaceae	0.6 (0.2)	6.3 (6.0)	0.2 (0.1)	0.1 (0.05)
f__Alcaligenaceae	0.6 (0,4)	0.1 (0.1)	0.2 (0.2)	0.3 (0.2)
f__Staphylococcaceae	0.5 (0.34)	13.9 (9.2)	18.3 (8.5)	24.1 (8.7)
f__Sphingobacteriaceae	0.4 (0.3)	0.6 (0.4)	0.2 (0.2)	0.4 (0.2)
f__Aerococcaceae	0.4 (0.4)	0.2 (0.2)	0.7 (0.4)	0.9 (0.6)
f__Prevotellaceae	0.3 (0.1)	0.1 (0.09)	0.4 (0.2)	0.1 (0.07)
f__Bifidobacteriaceae	0.2 (0.07)	0.02 (0.02)	0.2 (0.1)	0.1 (0.05)
f__Tissierellaceae	0.1 (0.1)	0.3 (0.2)	0.1 (0.1)	0.05 (0.05)
f__Peptostreptococcaceae	0.1 (0.1)	0.3 (0.1)	6.3 (4.0)	3.6 (3.3)
f__Phyllobacteriaceae	0.1 (0.1)	0.2 (0.1)	0.7 (0.5)	0.2 (0.1)
f__Porphyromonadaceae	0.4 (0.2)	0.1 (0.07)	0.4 (0.3)	0.3 (0.2)
f__Mogibacteriaceae	0.1 (0.04)	0.07 (0.06)	0.2 (0.1)	0.007 (0.007)
f__Veillonellaceae	0.1 (0.1)	0.1 (0.1)	0.1 (0.07)	0.2 (0.2)
f__Dermabacteraceae	0.06 (0.06)	0.3 (0.3)	1.3 (1.1)	3.0 (2.3)

Supplemental Table 2.2: Description of the unique OTUs detected in non-mastitic milk extracted from whole milk, fat, fat + pellet, and pellet.

	Mean Relative Abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Gemellaceae	0.6 (0.6)	0	0	0
f__Gordoniaceae	0.4 (0.4)	0	0	0
f__Deinococcaceae	0.4 (0.4)	0	0	0
f__Methylococcaceae	0.3 (0.3)	0	0	0
f__Campylobacteraceae	0.2 (0.1)	0	0	0
f__Aeromonadaceae	0.2 (0.1)	0	0	0
f__Brachyspiraceae	0.1 (0.1)	0	0	0
f__Sphingomonadaceae	0.08 (0.06)	0	0	0
f__Marinilabiaceae	0.04 (0.04)	0	0	0
f__Chthoniobacteraceae	0.03 (0.03)	0	0	0
f__Verrucomicrobiaceae	0.03 (0.02)	0	0	0
f__Halomonadaceae	0.02 (0.020	0	0	0
f__Erysipelotrichaceae	0.02 (0.01)	0	0	0
f__Fibrobacteraceae	0.01 (0.01)	0	0	0
f__RFP12	0.01 (0.01)	0	0	0
f__Thermoanaerobacteraceae	0.01 (0.01)	0	0	0
f__Thermoactinomycetaceae	0.008 (0.008)	0	0	0
f__Planctomycetaceae	0.007 (0.007)	0	0	0
f__Halobacteriaceae	0	3.6 (3.6)	0	0
f__Leptotrichiaceae	0	3.4 (3.4)	0	0
f__Rhizobiaceae	0	0.07 (0.07)	0	0
f__Hyphomicrobiaceae	0	0	1.1 (1.0)	0
f__Nocardiopsaceae	0	0	0.2 (0.2)	0
f__Hyphomonadaceae	0	0	0.1 (0.1)	0
f__Micromonosporaceae	0	0	0.1 (0.1)	0
f__Dehalobacteriaceae	0	0	0.05 (0.04)	0
f__Anaeroplasmataceae	0	0	0.04 (0.02)	0
f__Psychromonadaceae	0	0	0.02 (0.01)	0
f__Nocardiaceae	0	0	0	0.4 (0.2)
f__Glycomycetaceae	0	0	0	0.06 (0.04)
f__Rhabdochlamydiaceae	0	0	0	0.02 (0.02)
f__Thermobaculaceae	0	0	0	0.01 (0.01)

Supplemental Table 2.3: Description of the unique OTUs detected in *Escherchia coli* mastitic milk extracted from whole milk, fat, fat + pellet, and pellet.

	Mean Relative Abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Streptococcaceae	0.9 (0.9)	0	0	0
f__Bacillaceae	0.6 (0.6)	0	0	0
f__Micrococcaceae	0.1 (0.1)	0	0	0
f__Spirochaetaceae	0.02 (0.02)	0	0	0
f__Anaerolinaceae	0	0.9 (0.9)	0	0
f__Chthoniobacteraceae	0	0.6 (0.6)	0	0
f__Xanthomonadaceae	0	0.5 (0.5)	0	0
f__Leptotrichiaceae	0	0.5 (0.5)	0	0
f__Turicibacteraceae	0	0.2 (0.2)	0	0
f__Clostridiaceae	0	0.1 (0.1)	0	0
f__RFP12	0	0.1 (0.1)	0	0
f__Sphingomonadaceae	0	0.04 (0.04)	0	0
f__Halobacteriaceae	0	0.03 (0.03)	0	0
f__Flammeovirgaceae	0	0.01 (0.01)	0	0
f__Ulvophyceae	0	0	0.9 (0.9)	0
f__Peptostreptococcaceae	0	0	0	2.3 (2.3)
f__Bacteroidaceae	0	0	0	1.3 (0.8)
f__Moraxellaceae	0	0	0	0.5 (0.5)
f__Pirellulaceae	0	0	0	0.3 (0.5)
f__S24-7	0	0	0	0.2 (0.2)
f__Aerococcaceae	0	0	0	0.2 (0.2)
f__Veillonellaceae	0	0	0	0.1 (0.1)
f__Prevotellaceae	0	0	0	0.1 (0.1)
f__Dermacoccaceae	0	0	0	0.1 (0.1)
f__Paraprevotellaceae	0	0	0	0.05 (0.05)
f__Flavobacteriaceae	0	0	0	0.05 (0.05)
f__Porphyromonadaceae	0	0	0	0.05 (0.05)

Supplemental Table 2.4: Description of the unique OTUs detected in the microbial communities extracted from whole milk, fat, fat + pellet, and pellet of *Klebsiella* spp. mastitic milk.

	Mean Relative Abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Peptostreptococcaceae	2.5 (2.5)	0	0	0
f__Fusobacteriaceae	1.2 (1.2)	0	0	0
f__Lachnospiraceae	0.05 (0.05)	0	0	0
f__S24-7	0.04 (0.04)	0	0	0
f__Spirochaetaceae	0.04 (0.04)	0	0	0
f__Caulobacteraceae	0.03 (0.03)	0	0	0
f__Xanthomonadaceae	0.03 (0.03)	0	0	0
f__Bacillaceae	0	0.4 (0.4)	0	0
f__Porphyromonadaceae	0	0.03 (0.3)	0	0
f__Thermaceae	0	0.001 (0.001)	0	0
f__Desulfurococcaceae	0	0	0.04 (0.04)	0
f__Anaerolinaceae	0	0	0	0.3 (0.3)
f__Chlamydomonadaceae	0	0	0	0.04 (0.04)
f__Prevotellaceae	0	0	0	0.03 (0.03)
f__Pasteurellaceae	0	0	0	0.001580675

Supplemental Table 2.5: Description of the core microbiome, characterized as the OTUs at the family level detected in all *Streptococcus* spp. mastitic milk samples regardless of milk sample types (whole milk, fat, fat + pellet, and pellet).

	Mean Relative Abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Streptococcaceae	53.7 (8.8)	59.1 (9.0)	71.7 (9.0)	64.3 (9.4)
f__Pseudomonadaceae	4.0 (3.4)	5.5 (4.3)	0.12 (0.08)	3.3 (2.0)
f__Enterococcaceae	2.4 (2.4)	5.1 (3.61)	1.0 (1.0)	4.8 (4.8)
f__Ruminococcaceae	4.4 (3.1)	4.3 (4.2)	4.9 (4.7)	1.0 (0.6)
f__Flavobacteriaceae	3.6 (2.3)	4.0 (4.0)	0.6 (0.3)	0.2 (0.2)
f__Enterobacteriaceae	8.6 (4.0)	10.9 (5.8)	6.1 (4.5)	12.6 (6.0)
f__Brachyspiraceae	0.8 (0.8)	0.4 (0.4)	0.3 (0.3)	1.8 (1.8)
f__Desulfurococcaceae	0.4 (0.4)	0.3 (0.3)	2.5 (2.5)	1.7 (1.2)

Supplemental Table 2.6: Description of the unique OTUs detected in the microbial communities extracted from whole milk, fat, fat + pellet, and pellet of *Streptococcus* spp. mastitic milk.

	Mean Relative Abundance (Standard Error)			
Family	Whole Milk	Fat	Fat + Pellet	Pellet
f__Oxalobacteraceae	3.7 (3.7)	0	0	0
f__Bacillaceae	3.7 (3.0)	0	0	0
f__Aerococcaceae	1.1 (1.1)	0	0	0
f__Gordoniaceae	0.6 (0.5)	0	0	0
f__Halobacteroidaceae	0.4 (0.4)	0	0	0
f__Ulvophyceae	0.4 (0.4)	0	0	0
f__Caldicellulosiruptoraceae	0.4 (0.3)	0	0	0
f__Fusobacteriaceae	0.3 (0.3)	0	0	0
f__Deinococcaceae	0.2 (0.2)	0	0	0
f__Bogoriellaceae	0.2 (0.2)	0	0	0
f__Nocardioidaceae	0.2 (0.2)	0	0	0
f__Pseudonocardiaceae	0.2 (0.2)	0	0	0
f__Oceanospirillaceae	0.08 (0.8)	0	0	0
f__Phyllobacteriaceae	0.07 (0.07)	0	0	0
f__Staphylococcaceae	0.07 (0.07)	0	0	0
f__Xenococcaceae	0.06 (0.06)	0	0	0
f__Chitinophagaceae	0.03 (0.03)	0	0	0
f__Erythrobacteraceae	0.03 (0.02)	0	0	0
f__Pseudanabaenaceae	0.02 (0.02)	0	0	0
f__Streptomycetaceae	0.02 (0.02)	0	0	0
f__Geodermatophilaceae	0.01 (0.01)	0	0	0
f__Sphingomonadaceae	0.01 (0.01)	0	0	0
f__Tissierellaceae	0.01 (0.01)	0	0	0
f__Pasteurellaceae	0	0.1 (0.1)	0	0
f__Halomonadaceae	0	0.002 (0.002)	0	0
f__Sinobacteraceae	0	0	0.07 (0.074)	0
f__Bradyrhizobiaceae	0	0	0.06 (0.06)	0
f__Leptotrichiaceae	0	0	0.06 (0.06)	0
f__Mycoplasmataceae	0	0	0.06 (0.06)	0
f__Gemmataceae	0	0	0.02 (0.02)	0
f__Helicobacteraceae	0	0	0.02 (0.02)	0
f__Thermoanaerobacteraceae	0	0	0.005 (0.005)	0
f__Desulfohalobiaceae	0	0	0.004 (0.004)	0
f__Mogibacteriaceae	0	0	0.002 (0.002)	0
f__Comamonadaceae	0	0	0	2.0 (2.03)
f__Rhodobacteraceae	0	0	0	0.5 (0.5)
f__Corynebacteriaceae	0	0	0	0.1 (0.1)
f__Veillonellaceae	0	0	0	0.02 (0.02)
f__Prevotellaceae	0	0	0	0.01 (0.01)
f__S24-7	0	0	0	0.01 (0.01)

Supplemental Table 2.7: Description of the mean relative abundance of f__Staphylococcaceae among cow-health status groups (non-mastitic, *Escherichia coli*, *Klebsiella* spp. and *Streptococcus* spp.), milk fractions and DNA extraction kits.

	Mean Relative Abundance (Standard Error)	
f__Staphylococcaceae	PowerFood	PowerSoil
Non-mastitic		
Whole milk	0	0.9 (0.6)
Fat	13.9 (13.1)	13.9 (13.9)
Fat + Pellet	13.6 (10.1)	23.6 (14.5)
Pellet	20.4 (11.2)	28.6 (14.2)
Escherichia coli		
Whole milk	0	0
Fat	0	0
Fat + Pellet	1.7 (1.7)	1.6 (1.6)
Pellet	0.7 (0.7)	0.8 (0.8)
***Klebsiella* spp.**		
Whole milk	0	0
Fat	0	0
Fat + Pellet	0	0
Pellet	0	0
***Streptococcus* spp.**		
Whole milk	0	0
Fat	0	0
Fat + Pellet	0	0
Pellet	0	0

CHAPTER THREE

THE BOVINE COLOSTRUM MICROBIOME AND ITS ASSOCIATION WITH CLINICAL MASTITIS

Svetlana F. Lima[*], Andre G. V. Teixeira[*], Fabio S. Lima[†], Erika K. Ganda[*], Catherine H. Higgins[*], Georgios Oikonomou[‡]*, Rodrigo C. Bicalho*[1]

Journal of Dairy Science. 100: 3031–3042

*Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, New York, United States of America

[†]Department of Veterinary Medical Science, College of Veterinary Medicine, University of Illinois, Urbana, Illinois, United States of America

[‡]Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom

ABSTRACT

In an effort to characterize colostrum microbial diversity and its potential associations with early lactation clinical mastitis, we used high-throughput sequencing of the 16S rRNA gene to investigate the bovine colostrum microbiome. A prospective observational study was conducted that included 70 Holstein cows; colostrum samples were collected from all 4 mammary gland quarters. Colostrum samples were categorized according to whether the quarter was diagnosed (CMC) or not diagnosed (NCMC) with clinical mastitis during the first 30 days postpartum. Colostrum samples were dominated by Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Fusobacteria, and Tenericutes phyla, with the 6 most common taxa order (o), family (f), and genus (g) being *g__Staphylococcus, g__Prevotella, f__Ruminococcaceae, o__Bacteroidales, o__Clostridiales,* and *g__Pseudomonas*. The colostrum microbiota of primiparous cows was significantly richer (higher number of bacterial species) than that of multiparous cows, and differences in colostrum taxonomic structure between parities were also observed. The microbial community of NCMC samples of primiparous cows was significantly more diverse than that of CMC samples, and the relative abundances of the Tenericutes and Fusobacteria phyla as well as the *Mycoplasma* and *Fusobacterium* genera were significantly higher in NCMC than in CMC samples of primiparous cows. The colostrum core microbiome, defined as the bacterial taxa common to all colostrum samples examined, was composed of 20 taxa and included bacterial genera already known to be associated with mastitis (e.g. *Staphylococcus*, *Mycoplasma*, and *Streptococcus* spp.). Our results indicate that the colostrum microbiome of primiparous cows differs from that of multiparous cows, and it harbors some diversity and taxonomic markers of mammary gland health specific to primiparous cows only.

Keywords: colostrum, mastitis, microbiome

INTRODUCTION

Colostrum is a complex biological fluid that supports the growth and health of the neonate (Ogra et al., 1977, Chase et al., 2008, Hammon et al., 2013), and plays a key role during this initial and critical period of life. Bovine colostrum (BC) is not only nutrient rich, it also contains growth factors and antimicrobial agents (Pakkanen et al., 1997, Elfstrand et al., 2002), and harbors apparently beneficial microflora such as members of the *Bifidobacterium* and *Lactobacillus* genera (Lindner et al., 2011) that are widely used as probiotics. Culture-independent microbiome studies are greatly improving our understanding of the microbial diversity of different bodily environments, such as human or bovine milk (Cabrera-Rubio et al., 2012, Kuehn et al., 2013, Ward et al., 2013, Oikonomou et al., 2014). Indeed, studies in humans have described the microbial composition of breast milk during lactation (Cabrera-Rubio et al., 2012, Ward et al., 2013, Khodayar-Pardo et al., 2014) and its potential effects on infant health (Ward et al., 2013). Human colostrum in these studies was also shown to harbor a diverse microbial community that was different from that of milk. Our group previously characterized the bovine milk microbial community and its relationship with mammary gland health (Oikonomou et al., 2012, Oikonomou et al., 2014). However, only a few studies have investigated the microbial composition of bovine (Lindner et al., 2011) and human (Cabrera-Rubio et al., 2012) colostrum. Lindner (2011) isolated and sequenced 29 bacterial strains from BC samples; however, culture-independent characterization of the BC microbiome has not yet been conducted, and its potential impact on mammary gland health remains unknown.

In dairy cattle, colostrum is the first milk produced after a non-lactating period, commonly referred to as the dry period (Watters et al., 2008). The dry period is critical for replacement of senescent mammary epithelial cells as well as the elements produced by those cells (Capuco et al., 1997). Studies have shown an increased risk of mammary gland infection correlated with intra-mammary infection (**IMI**) at the beginning and/or end of the dry period (Aarestrup and Jensen et al., 1997, Green et al., 2002, Bradley and Green et al., 2004). Furthermore, DNA fingerprinting analysis demonstrated that the pathogens which cause infection in the dry period may persist, and that the infected udder quarter is more likely than the uninfected quarters to develop mastitis as a consequence of these dry-period pathogens (Bradley and Green et al., 2000). Because significant numbers of IMIs are detected in both the dry period and the early fresh period, targeting new intervention procedures to these periods may help to improve udder health and milk production (Natzke et al., 1981). Thus, characterization of the bovine colostrum microbial community could help identify the bacterial profiles closely associated with mammary gland infection soon after parturition.

Inflammation of the mammary gland parenchyma (mastitis) is mainly triggered by bacterial infection (Federation, 1987, Oviedo-Boyso et al., 2007). Bovine mastitis resulting from bacterial infection is a huge economic burden for the dairy industry worldwide. The estimated cost of each case of clinical mastitis (CM) which occurs in the first 30 days of lactation is approximately $444, which includes direct costs such as therapeutics, non-saleable milk, and cow death, and indirect costs such as future milk production loss and premature culling (Rollin et al., 2015). The United States Department of Agriculture National Animal Health Monitoring Service (NAHMS) reported in 2007 that CM affects 94.9% of American dairy herds (USDA,

2007). Additionally, the reported incidence of CM per American farm unit during the same study period was approximately 16.5% (USDA, 2007).

Thus, the aim of this study was to use metagenomic high-throughput sequencing of the 16S rRNA gene to assess the microbial community of colostrum of Holstein dairy cows and evaluate its potential association with early CM.

MATERIALS AND METHODS

Ethics statement

The research protocol was reviewed and approved by the Cornell University Institutional Animal Care and Use Committee (protocol number 2013-0056). The methods were carried out in accordance with the approved guidelines.

Study design and study population

We used a prospective observational cohort study design in which colostrum samples were aseptically collected daily (morning and afternoon) at the first milking within 8 hours after parturition on a large commercial dairy farm situated in upstate New York. After colostrum sample collection, the cohort of animals was followed for 30 days postpartum for mammary gland health characterization. Colostrum samples were obtained from each mammary gland quarter defined as: left front (LF), left hind (LH), right front (RF), and right hind (RH) cow udder quarters. All post-partum cows that were first milked within 8 hours after parturition from November to December of 2013 were eligible to be enrolled in this study, and a total of 100 Holstein dairy cows were sampled at first milking. After sample exclusion due to the following reasons — absence of cow health report; missed colostrum sample collection due to udder-

quarter trauma or structural disorder; problems with downstream screening (e.g. failed DNA amplification or poor sequencing quality) — a total of 280 samples from 70 cows was submitted for microbiome analysis. In the present study all 280 colostrum samples presented the same physical appearance (yellow hue color, high fat content, and no sign of inflammation or trauma such as blood, flakes, or clot).

Animals and facilities

The commercial dairy farm involved in the study milked 3,450 Holstein cows thrice daily in a double 52-stall parallel milking parlor. During the study period, the average milk production per cow per day was 40 kg (standard error (SE): 4.41), and the incidence of mastitis was 19.2%. Following the farm's protocol, lactating pregnant cows were submitted to a dry period of 60 days, and at dry off, multiparous cows received intramammary antibiotic therapy (Orbenin®-DC, Merck Animal Health). Pregnant heifers and cows were housed together in free-stall barns with concrete stalls covered with rubber mattresses and bedded with dried manure solids, separated from the lactating cows. Pregnant animals were housed into two separate groups: the "far-off" group (where dry cows remained until two weeks before expected calving), and the "close-up" group (cows in their last 2 weeks before expected calving). Pregnant animals that were at stage 1 or 2 of parturition were moved into the deep-bedded maternity barn, with four identical group-pens (a total of 3 cows per pen), to calve. Right after parturition, calves were removed from the maternity pen and placed into a newborn pen bedded with dry sawdust and heated with heating lamps. Fresh cows were first milked within 8 hours of calving in the double 52-stall parallel milking parlor.

Prepartum heifers and cows were given a high-fiber, low energy density diet, whereas postpartum cows were given a low-fiber content, high energy density diet derived from higher starch and fat supplementation. The diet was formulated to meet or exceed the nutrient requirements for lactating Holstein cows weighing 650 kg and producing 45 kg of 3.5% fat corrected milk (NRC, 2001).

Sample collection

One member of our research group collected individual colostrum quarter samples aseptically at the first milking, within 24 hours of parturition. Sampling methods followed standard recommendations by the National Mastitis Council's Laboratory Handbook on Bovine Mastitis (NM, 1999). The first streams of colostrum from each quarter were discarded for mammary gland stimulation, and subsequently the teats were dipped in iodine tincture. The teats were then exposed to the disinfectant for 30 seconds and dried by using an individual towel, as performed by the farm personnel. Research personnel performed the physical scrubbing of the teats with gauze soaked in 70% alcohol. Once more, the first streams of colostrum were discarded and the colostrum samples were then collected. A 10-ml aliquot of colostrum was collected in a sterile 50-ml centrifuge tube (VWR International, Radnor, PA). Samples were kept on ice, transported to the laboratory and stored at -20 °C.

Case definition

Clinical mastitis examination was performed at the milking parlor, thrice daily on all lactating cows, by farm employees with extensive experience and trained by Cornell University veterinarians (Ambulatory and Production Medicine Department). Fore-stripping was performed

and quarters presenting abnormal milk (i.e. presence of flakes, clots, or serous milk), independently of systemic illness and signs of inflammation of the mammary gland, were diagnosed as CM.

DNA extraction

Isolation of DNA from colostrum samples was performed by adapting the protocol of Rasolofo (2010) (Rasolofo et al., 2010). A 10-ml aliquot of colostrum was clarified by the addition of 2 ml of 0.5 M EDTA (ethylenediamine tetraacetic acid) at pH 8.0, followed by centrifugation (9000 rpm at 4°C for 30 min). The supernatant containing fat and protein was discarded and the cell pellet was washed twice with 1 ml of sucrose buffer (12% sucrose, 25 mM Tris-HCl, pH 8.0) followed by centrifugation (9000 rpm at 4°C for 30 min). The pellet was re-suspended in 0.4 ml of sucrose buffer containing 800 µg of lysozyme and incubated for 1 hour at 37°C. Finally, total metagenomic DNA was extracted by using a Powersoil® DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA) following the manufacturer's protocol. The DNA concentration was evaluated by optical density using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 230, 260 and 280 nm.

PCR amplification of the V4 hypervariable region of the bacterial 16S rRNA gene

Amplification of the 16S rRNA gene was performed by PCR using barcoded primers on all metagenomic DNA samples of colostrum. For amplification of the V4 hypervariable region of the bacterial/archaeal 16S rRNA gene, primers 515F and 806R were used according to previously described methods and optimized for the Illumina MiSeq platform (Caporaso et al., 2012). The Earth Microbiome Project (http://www.earthmicrobiome.org/) (Gilbert et al., 2010) was used to select 140 different 12-bp error-correcting Golay barcodes for the procedure, as

previously described (Caporaso et al., 2012). The 5'-barcoded amplicons were generated in triplicate using 12-300 ng of template DNA, 2X EconoTaq® Plus Green Master Mix (Lucigen®, Middleton, WI) and 10 µM of each primer. The PCR conditions for amplification of the 16S rRNA gene included an initial denaturing step of 94°C for 3 min followed by 35 cycles of 94°C for 45 s, 50°C for 1 min and 72°C for 90 s and a final elongation step of 72°C for 10 min. Replicate amplicons were pooled and purified using the Gel PCR DNA Fragment Extraction kit (IBI Scientific, Peosta, IA) and visualized by electrophoresis through 1.2% (wt/vol) agarose gels stained with 0.5 mg/ml ethidium bromide. Blank controls in which no DNA was added to the reaction were performed, and no amplification was detected. Purified amplicon DNA was quantified using the Quant-iT™ PicoGreen® and dsDNA Broad Range Assay Kit (Life Technologies Corporation, Carlsbad, CA).

Sequence library analysis

Aliquots of colostrum amplicon samples were diluted to the same concentration and then pooled into three different runs according to individual barcode primers for the 16S rRNA gene. Final equimolar libraries were sequenced using the MiSeq reagent kit v2 (300 cycles) on the MiSeq platform (Illumina Inc., San Diego, CA). The 16S rRNA gene sequences generated were processed through the open source software pipeline Quantitative Insights Into Microbial Ecology (QIIME) version 1.7.0-dev (Caporaso et al., 2012). Sequences were filtered for quality using established guidelines (Bokulich et al., 2013). Sequences were binned into Operational Taxonomic Units (OTU) based on 97% identity using UCLUST (Edgar et al., 2010) against the Greengenes reference database (McDonald et al., 2012), May 2013 release. Low-abundance clusters were filtered and chimeric sequences were removed using USEARCH (Edgar, 2010).

The representative sequences for each OTU were compared against the Greengenes database for taxonomy assignment, and only full-length, high-quality reads (-r = 0) were used for analysis. The output of this workflow is a classification of reads at multiple taxonomic levels: kingdom (k), phylum (p), class (c), order (o), family (f), genus (g) and species (s).

Shannon diversity and Chao1 indexes were calculated using the QIIME pipeline. Before estimating the Shannon and Chao1 indexes, all sample libraries were rarefied to an equal depth of 10,000 sequences using QIIME.

Sample categorization and statistical analysis

Colostrum samples from mammary gland quarters that were diagnosed with CM in the first 30 days postpartum were categorized as clinical mastitic colostrum (CMC), and those samples from mammary gland quarters that were free of CM during the first 30 days postpartum were categorized as non-clinical mastitic colostrum (NCMC).

Descriptive statistics for the total number of cows, quarters, and CM incidence (after applying the exclusion criteria) were assessed using the Distribution platform offered by JMP Pro 11 (SAS Institute INC). The Pearson chi-square test was used to compare the incidence of mastitis between parities.

Sample richness was calculated by using the Chao 1 richness index, and diversity by using Shannon diversity index. The Chao 1 index is a nonparametric procedure that estimates the minimum richness (number of OTU) in a sample by using the number of rare OTU (singletons and doublets) (Chao et al., 1984). The Shannon index is calculated based on both richness and evenness (how evenly the taxa are distributed in a sample) (Zand et al., 1976). The Shannon index increases both when the number of species and evenness increases. To evaluate differences

of the Chao 1 and Shannon indexes between the groups tested (primiparous vs. multiparous, and CMC vs. NCMC), general linear models was fitted in JMP Pro 11. Two models were created: one with parity being the only independent variable (and which used colostrum samples only from quarters that did not experience CM during the first 30 days portpartum), and the other, in which colostrum category (NCMC and CMC) blocked by parity was offered to the model as the only independent variable. The number of reads in each group tested (parity, colostrum category) was determined, and a similar procedure to the one described above was conducted.

Taxa data and relative abundance (proportion of each bacterial type within the total number of sequences in the sample) were used to describe the most abundant phyla and other OTU levels of the colostrum samples. To evaluate differences between microbial taxa of colostrum from primiparous and multiparous cows, colostrum samples only from quarters that did not experience CM during the first 30 days postpartum were used. A general linear model was fitted in JPM Pro 11 (SAS Institute Inc., NC). Parity, udder quarter, and the interaction between these terms were used as independent variables. Least squares mean relative abundance and standard error of the mean for each bacterium were obtained.

Differences between colostrum categories (NCMC and CMC) were also evaluated using general linear models fitted in JPM Pro 11 (SAS Institute Inc., NC) software. Microbial relative abundance (phyla and other OTU levels) in NCMC and CMC, blocked by parity, were compared, and the least squares mean relative abundance and standard error of the mean for each bacterium were obtained. A more detailed analysis of the differences in the relative abundances of *g__Staphylococcus*, *g__Pseudomonas*, *g__Fusobacterium*, *g__Mycoplasma*, *g__Corynebacterium*, *g__Streptococcus*, and *g__Escherichia* between colostrum categories, blocked by parity, was conducted using general linear models fitted in JPM Pro 11.

The core microbiome was defined as all taxa found to be present across all colostrum samples. A general linear model was performed and fitted in JMP Pro 11 to describe the least squares mean relative abundance of the taxa shared by all colostrum samples of all primiparous and multiparous cows. The independent variable offered to the model was udder quarter category, and blocked by parity.

RESULTS

Sequencing results, descriptive statistics, and alpha diversity indexes

Sequences were filtered for size, quality, and for the presence of chimeras. The total post-quality-control number of sequences used in the study was 21,158,985. The average coverage was 75,568 sequences per sample, with a standard deviation (SD) of 22,901 and a range of 7,517 to 151,326. The number of sequences evaluated did not differ among parities, and colostrum categories (Supplemental 3.1).

Descriptive statistics for the number of cows, number of quarters, and CM incidence are presented in Table 3.1. After applying the cow exclusion criteria, the incidence of CM at cow level for primiparous cows was 46.2% (n=12) and for multiparous 22.7% (n=10).

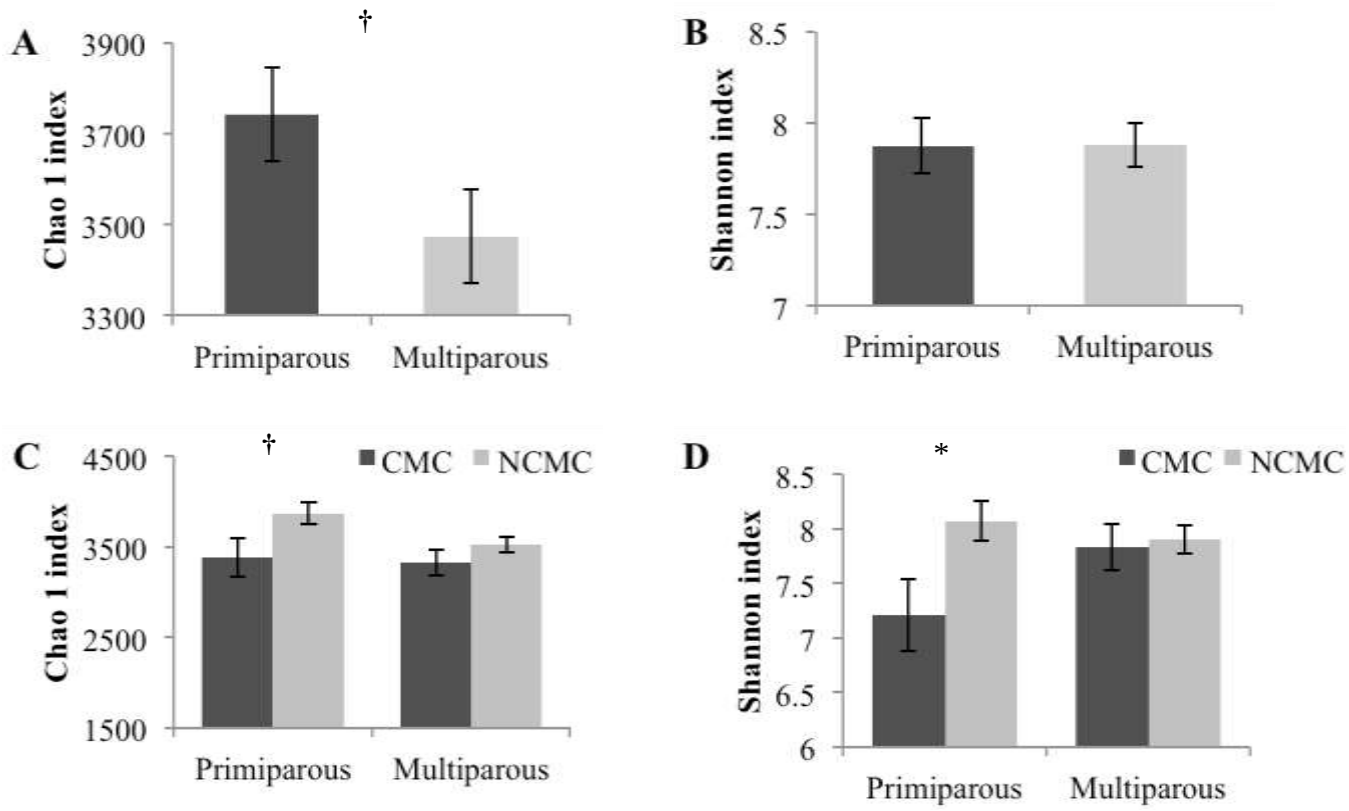
The Chao 1 richness and Shannon diversity indexes were calculated to obtain estimates of community diversity (Figures 3.1A, B, C & D). The Chao 1 index was higher in primiparous than in multiparous cows ($P < 0.05$, Figure 3.1A), and tended to be higher in NCMC than in CMC samples of primiparous cows ($P = 0.054$, Figure 3.1C). The Shannon index was also higher for NCMC when compared to CMC samples of primiparous cows ($P < 0.05$, Figure 3.1D). The Shannon index was not significantly different when parities were compared, nor when the colostrum categories of multiparous cows were compared (Figure 3.1B & D,

respectively). Similarly, the Chao 1 index did not differ significantly between the colostrum categories for multiparous cows (Figure 3.1C).

Table 3.1. Incidence of clinical mastitis during the first 30 days postpartum by cow and quarter for primiparous and multiparous cows used in the study. Numbers in brackets indicate the number of quarters with clinical mastitis in each category, and *P*-value represents the difference between parities.

	Primiparous	**Multiparous**	***P*-value**
Total number of cows	26	44	-
Total number of quarters	104	176	-
Incidence of mastitis by quarter	23.1% (24)	29.0% (51)	0.003
Left front	25.0% (6)	21.6% (11)	0.85
Left hind	20.8% (5)	27.5% (14)	0.25
Right front	25.0% (6)	27.5% (14)	0.43
Right hind	29.2% (7)	23.5% (12)	0.97

Figure 3.1: Chao 1 richness index and Shannon diversity index according to parity (A & B, respectively) and colostrum category within parity (C & D, respectively). Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$, and * represents $P < 0.05$.

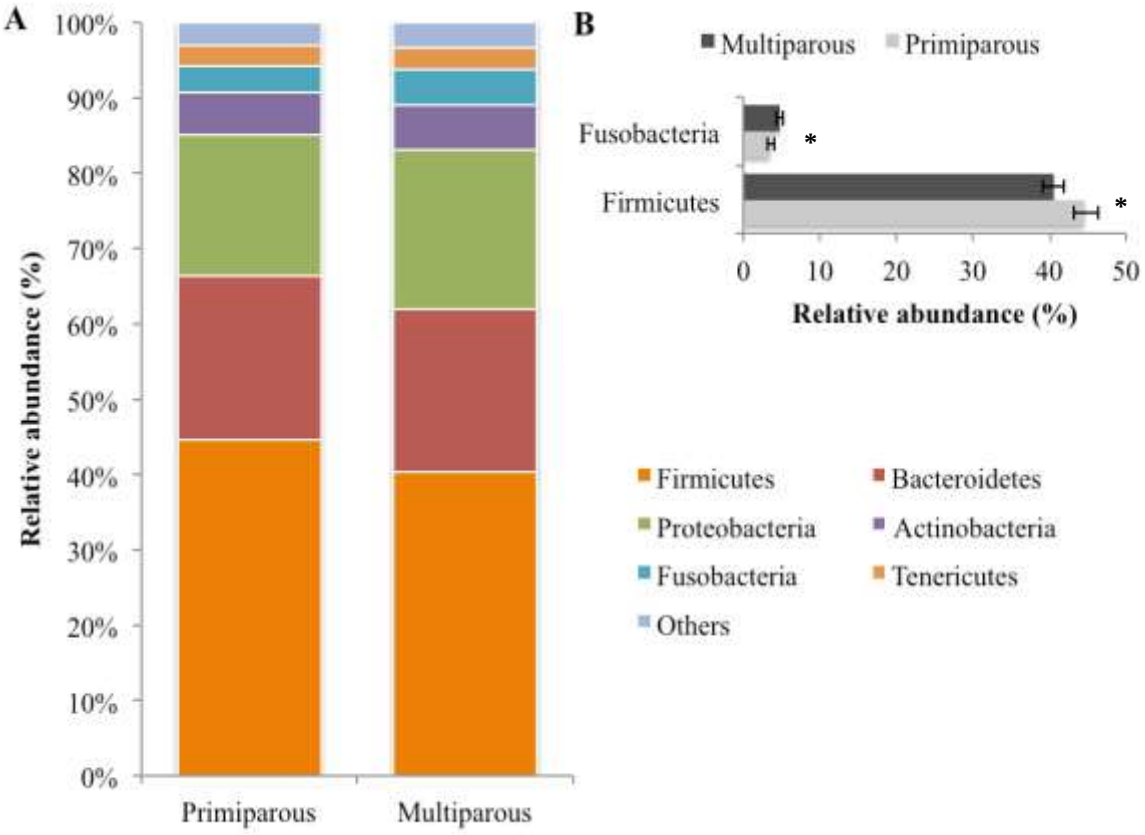


Phylogenetic profile and core microbiome

Alignment of OTU at the 97% similarity threshold against the Greengenes database enabled identification of the colostrum bacterial community at the phylum and genus levels. Whereas a majority of OTU was classified at the g level, some were classified only at the p, c, o, or f level. The six most common bacterial phyla detected in colostrum samples from quarters that did not experience CM during the first 30 days postpartum are depicted in Figure 3.2A. Of the 6 most common phyla detected in both parities, the mean relative abundance (MRA) of the Firmicutes phylum was higher ($P < 0.05$) in samples of primiparous (MRA: 46.1%; standard

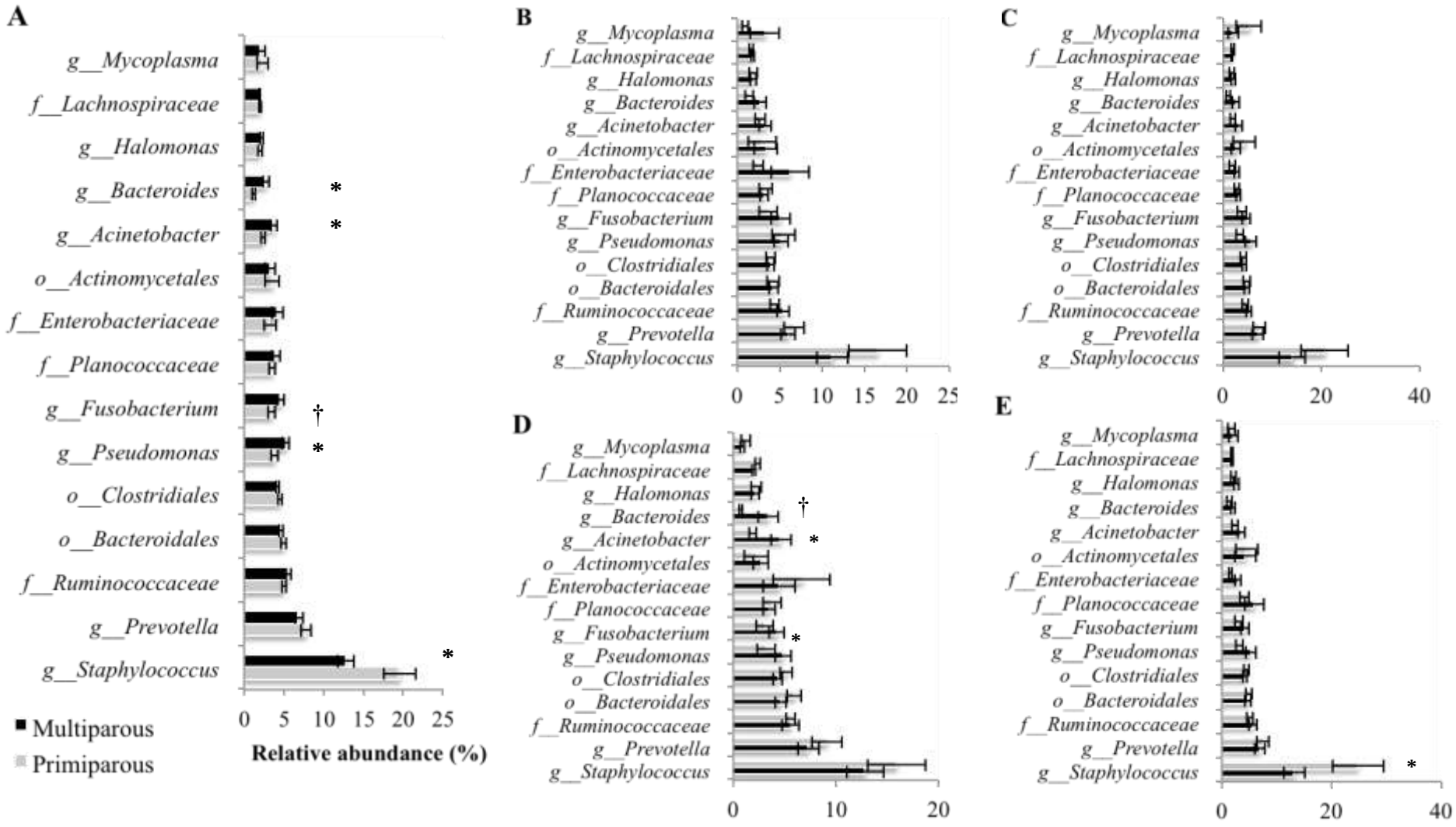
error (**SE**): 1.8) than in multiparous cows (MRA: 40.8%; SE: 1.6), and the Fusobacteria phylum were lower ($P < 0.05$) in samples of primiparous (MRA: 2.8%; SE: 0.6) than in multiparous cows (MRA: 4.8%; SE: 0.5) (**Figure 3.2B**).

Figure 3.2: Aggregate microbiome composition at the phylum level for 16S rRNA sequences of colostrum samples (A), and the mean relative abundance of Fusobacterium and Firmicutes phyla according to parity (B). Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$, and * represents $P < 0.05$.



The 15 most common OTU (o, f, and g) detected in colostrum samples of primiparous and multiparous cows are shown in Figure 3.3A, and more detailed findings at the quarter level are presented in Figure 3B-E. *g__Bacteroides* (MRA: 1.1%, SE: 0.4 for primiparous and MRA: 2.5%, SE: 0.3 for multiparous, $P < 0.05$), *g__Acinetobacter* (MRA: 2.3%, SE: 0.5 for primiparous and MRA: 3.5%, SE: 0.4 for multiparous, $P < 0.05$), *g__Pseudomonas* (MRA: 3.5%, SE: 0.5 for primiparous and MRA: 4.8%, SE: 0.4 for multiparous, $P < 0.05$), and *g_Fusobacterium* (MRA: 2.8%, SE: 0.6 for primiparous and MRA: 5.0%, SE: 0.5 for multiparous, $P = 0.08$) were either significantly higher or tended to be significantly higher in colostrum samples of multiparous cows than in primiparous cows (Figure 3.3A). Colostrum samples from primiparous cows had a higher abundance ($P < 0.05$) of *g__Staphylococcus* (MRA: 19.4%, SE: 1.9 for primiparous and MRA: 13.6%, SE: 1.5 for multiparous, $P < 0.05$) when compared to multiparous cows (Figure 3.3A). At the cow udder quarter level, microbial abundance differences between the parity samples were observed only in the RF and RH quarters (Figures 3.3D & E).

Figure 3.3: Bar graphs illustrating the 15 most common microbial taxa according to parity (primiparous & multiparous, A) and according to left front (B), left hind (C), right front (D), and right hind (E) cow udder quarters and parity. *o* (order), *f* (family), and *g* (genus) represent the taxonomic level of the described bacterium. The mean relative abundance is represented by x axis values. Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$, and * represents $P < 0.05$.



Microbial abundance differences at the phylum (Figures 3.4A & B) and other OTU levels (o, f, g) (Figures 3.5 & 6) between the colostrum categories (NCMC and CMC) were assessed. Among the most common phyla detected in the NCMC and CMC samples of primiparous and multiparous cows (Figure 3.4B), the relative abundances of the Tenericutes (MRA: 5.6%, SE: 1.4 for CMC, and MRA: 2.1%, SE: 0.7 for NCMC) and Fusobacteria (MRA: 5.8%, SE: 0.9 for CMC, and MRA: 2.8%, SE: 0.5 for NCMC) phyla were higher ($P < 0.05$) in the CMC than in the NCMC samples of primiparous cows (Figure 3.4B). Descriptive cataloging of the 15 most common bacterial taxa (c, o, f, and g) detected in colostrum samples of primiparous and multiparous cows according to each colostrum category and quarter evaluated is presented in Supplemental 3.2A & B. The genera g__*Staphylococcus*, g__*Pseudomonas*, *g__Fusobacterium*, *g__Mycoplasma*, *g__Corynebacterium*, *g__Streptococcus*, and *g__Escherichia* were subjected to more detailed analyses (Figure 3.5A & B). *g__Fusobacterium* and *g__Mycoplasma* were relatively more abundant ($P < 0.05$) in CMC than in NCMC samples of primiparous cows (Figure 3.5A), and *g__Streptococcus* tended to be more abundant in CMC than in NCMC samples of multiparous cows (Figure 3.5B).

The colostrum core microbiome, defined as the set of microbial organisms persistently present in all samples evaluated, regardless of parity or colostrum category, was assessed. In total, 20 bacterial taxa were detected in all colostrum samples from all quarters of all cows examined (Table 3.2).

Figure 3.4: Aggregate microbiome composition at the phylum level for 16S rRNA gene sequences of colostrum samples according to parity (primiparous & multiparous), and colostrum category (non-clinical mastitic colostrum & clinical mastitic colostrum) (A). The mean relative abundance of the Tenericutes and Fusobacteria phyla in primiparous cows according to colostrum category (B). Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$, and * represents $P < 0.05$.

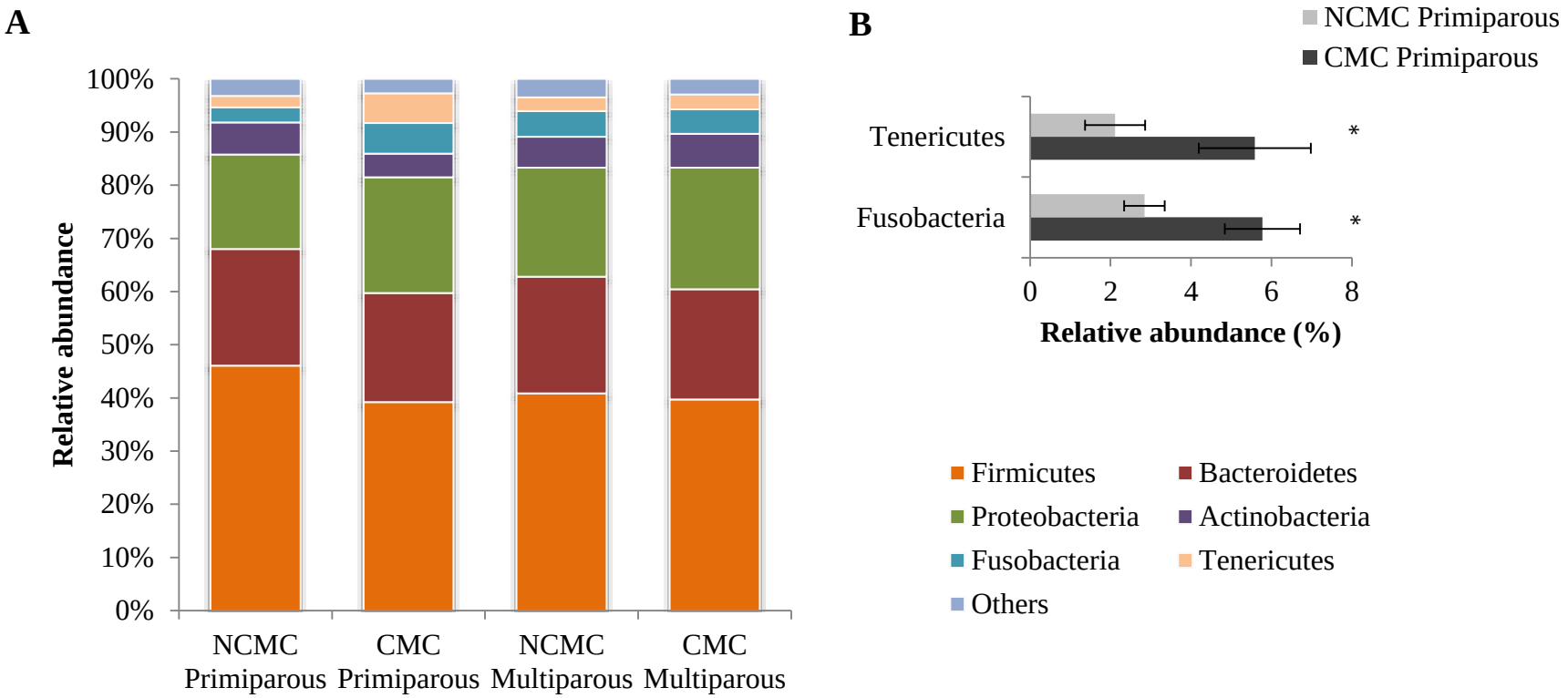


Figure 3.5: Bar graphs illustrating the mean relative abundance of *g_Staphylococcus*, *g_Pseudomonas*, *g_Fusobacterium*, *g_Mycoplasma*, *g_Corynebacterium*, *g_Streptococcus*, and *g_Escherichia* detected in the clinical mastitic colostrum category (black) and the non-clinical mastitic colostrum category (gray) of primiparous (A) and multiparous (B) cows. *o* (order)*, f* (family), and *g* (genus) represent the taxonomic level of the described bacterium. Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$, and * represents $P < 0.05$.

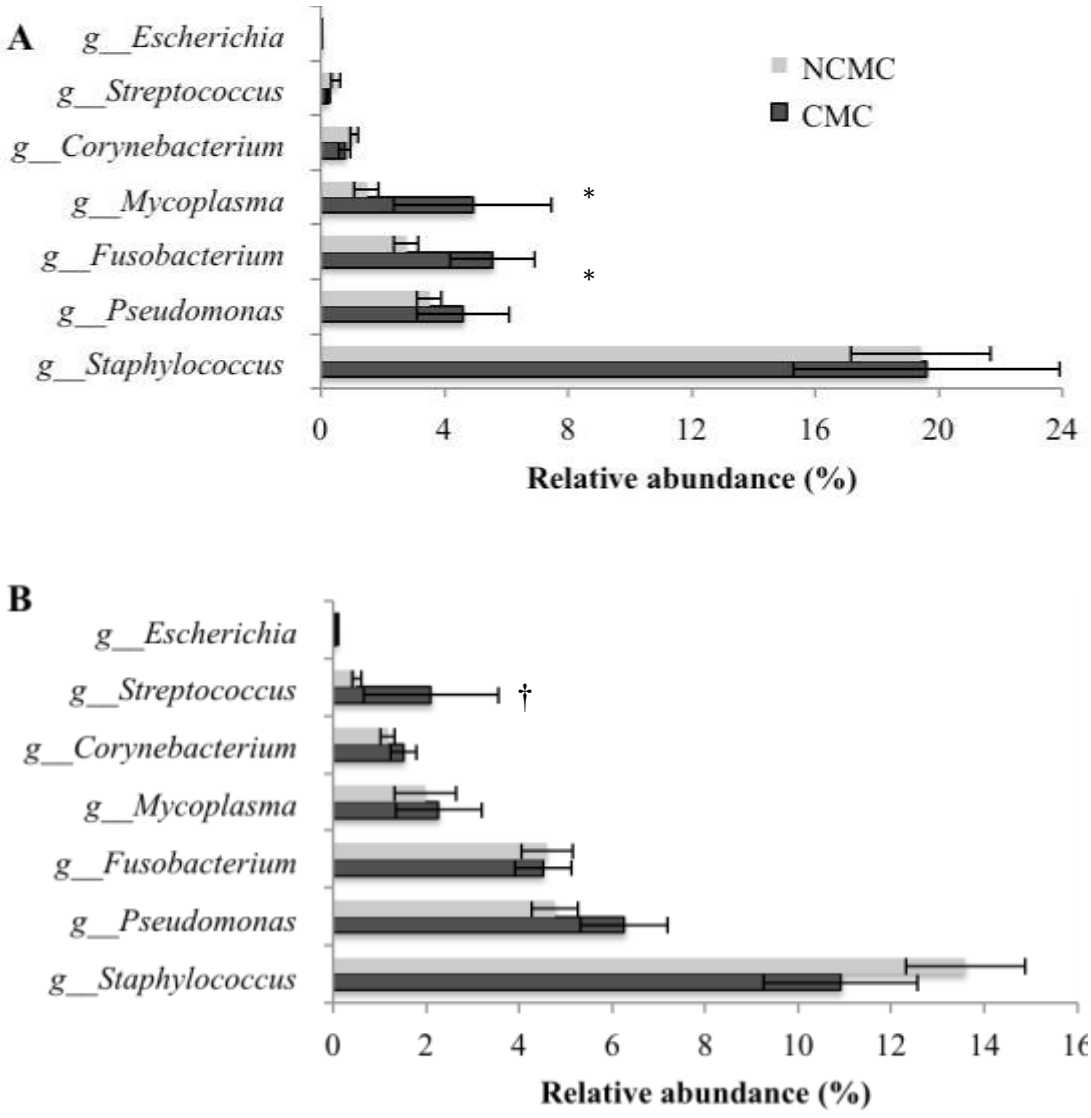


Table 3.2. Relative abundances of colostrum core microbiome taxa, defined as the bacterial taxa detected in all colostrum samples from all quarters of all cows evaluated, according to parity (multiparous & primiparous), and udder quarter (left front, left hind, right front & right hind). *o* (order*), f* (family), and *g* (genus) represent the taxonomic level of the described bacterium.

	Primiparous								Multiparous							
OTU	[1]LF (%)	[2]SE	[3]LH (%)	[2]SE	[4]RF (%)	[2]SE	[5]RH (%)	[2]SE	[1]LF (%)	[2]SE	[3]LH (%)	[2]SE	[4]RF (%)	[2]SE	[5]RH (%)	[2]SE
g__Staphylococcus	16.5	3.4	20.7	4.7	15.9	2.8	24.8	4.6	11.2	1.8	14.1	2.6	12.8	1.8	13.1	1.9
g__Prevotella	6.7	1.2	7.5	1.3	9.1	1.5	7.5	1.0	6.0	0.8	7.1	1.1	7.3	1.0	6.9	0.9
f__Ruminococcaceae	4.4	0.6	4.6	0.6	5.6	0.5	5.0	0.5	5.4	0.7	5.4	0.6	5.6	0.8	5.6	0.7
o__Bacteroidales	4.2	0.7	4.8	0.7	5.9	0.7	4.9	0.5	4.3	0.5	4.8	0.6	4.6	0.5	4.7	0.5
o__Clostridiales	3.9	0.6	4.2	0.6	5.2	0.6	4.5	0.5	4.0	0.4	4.4	0.4	4.3	0.4	4.2	0.4
g__Pseudomonas	5.4	1.3	3.3	0.7	3.2	0.8	3.1	0.6	5.1	0.8	5.7	1.1	4.8	0.8	5.2	0.9
g__Fusobacterium	3.6	1.1	3.9	0.9	3.1	0.8	3.1	0.7	5.1	1.1	4.8	0.8	4.2	0.7	4.2	0.8
f__Planococcaceae	3.3	0.8	2.8	0.6	3.8	0.9	4.1	0.9	3.2	0.5	3.3	0.6	3.5	0.6	5.9	1.7
f__Enterobacteriaceae	2.5	0.6	1.9	0.6	6.7	2.8	1.6	0.3	6.2	2.3	2.8	0.5	4.5	1.5	2.9	0.5
o__Actinomycetales	2.9	1.6	4.4	2.2	2.2	1.2	4.5	2.1	3.4	1.4	2.6	1.0	2.7	0.8	4.2	1.9
g__Acinetobacter	2.8	0.6	2.1	0.5	1.9	0.3	2.5	0.6	3.3	0.7	3.2	0.7	4.6	1.0	3.5	0.7
g__Bacteroides	1.4	0.4	1.1	0.5	0.7	0.2	1.4	0.4	2.7	0.7	2.6	0.7	3.4	1.0	2.0	0.5
g__Halomonas	1.9	0.5	1.8	0.5	2.3	0.6	2.1	0.5	1.9	0.3	2.1	0.4	2.2	0.4	2.6	0.5
f__Lachnospiraceae	1.7	0.3	2.0	0.3	2.3	0.2	1.8	0.2	1.8	0.2	1.8	0.2	2.0	0.2	1.8	0.2
g__Mycoplasma	0.9	0.4	5.3	2.5	1.2	0.4	1.8	0.6	3.3	1.7	2.0	1.0	0.9	0.2	2.1	0.8
g__Ruminococcus	1.4	0.2	1.5	0.2	1.7	0.2	1.6	0.2	1.4	0.1	1.5	0.2	1.4	0.2	1.4	0.1
g__Corynebacterium	1.1	0.2	0.9	0.2	1.2	0.2	0.8	0.2	1.1	0.2	1.5	0.4	1.1	0.2	1.5	0.3
f__Pseudomonadaceae	2.2	1.0	1.0	0.2	1.0	0.2	1.0	0.2	1.9	0.6	1.2	0.3	1.4	0.3	1.1	0.2
g__Porphyromonas	1.0	0.3	1.7	0.7	0.7	0.2	0.9	0.3	1.1	0.3	1.5	0.4	0.8	0.2	1.3	0.3
g__Bacillus	1.0	0.4	0.6	0.2	0.8	0.2	0.6	0.1	2.1	1.3	0.7	0.1	1.0	0.3	0.8	0.1

[1]Left front, [2]Left hind, [3]Right front, [4]Right hind, [5]Standard error

DISCUSSION

To our knowledge, this is the first detailed description of the bovine colostrum microbiome. Colostrum is the first milk following a long non-lactating period known as the dry period. The dry period is an interval of critical susceptibility to mammary gland infections (Green et al., 2002, Bradley and Green et al., 2004). Therefore, characterization of the BC microbiota may improve our understanding of the importance of the microbial community profile in early mastitis development. The results of our study indicate that colostrum harbors a rich and diverse microbial community, regardless of parity (primiparous or multiparous) or early lactation infection status (CMC or NCMC). Furthermore, we found differences in the bacterial taxonomic structure of colostrum between primiparous and multiparous cows, as well as between NCMC and CMC of primiparous cows.

Another interesting finding of our study is that the colostrum microbiota of primiparous cows was richer than that of multiparous cows. A possible explanation for this result is that these two groups are physiologically different; most obviously, the multiparous cows experienced at least one previous lactation, potentially including IMI events (Falentin et al., 2016). The use of dry-period antibiotic therapy in multiparous cows is another plausible explanation for these findings. Although the intramammary antibiotic used at drying-off in our study had a half-life of 4 weeks, it could still have a lasting effect on the mammary gland and ultimately on the colostrum microbiome.

The OTU *g__Staphylococcus, g__Prevotella, f__Ruminococcaceae, 0__Bacteroidales, 0__Clostridiales,* and *g__Pseudomonas* were the 6 dominant taxa we identified in colostrum samples from primiparous and multiparous cows. In a recently published review, Addis (2016) described the potential existence of an endogenous entero-mammary pathway, in which intestinal

bacteria “migrate” to the mammary gland, and this could explain the presence of gut bacteria such as *g__Prevotella,* and *f__Ruminococcaceae* in colostrum samples (Addis et al., 2016). Furthermore, a limited microbiota of BC was previously described by Lindner (2011), in which 29 strains, including *Lactococcus casei*, *Staphylococcus pseudintermedius*, *Staphylococcus chromogenes*, *Bifidobacterium pseudolongum*, and *Propionibacterium acnes*, were isolated from colostrum and identified by their 16S rRNA gene sequences (Lindner et al., 2011). The distinct bacterial organisms found by our group may be attributed in part to the methodology and technology we used. Whereas Lindner (2011) extracted DNA from culture isolates followed by sequencing of the 16S rRNA gene, we opted for a culture-independent approach. Traditionally, the bacterial ecology of milk has been described by culture-dependent methods, which are time-consuming and limited in terms of their ability to detect a truly representative microbial community (Taponen et al., 2009, Oikonomou et al., 2012). In contrast, 16S rRNA gene sequencing is a much more robust approach for determination of the complete bacterial community in an environment as well as taxon relative abundances (Amann et al., 1995, Schloss and Handelsman et al., 2005).

g__Staphylococcus was the most abundant genus identified in colostrum samples, independent of quarter, parity, and mastitis status. Furthermore, *g_Staphylococcus, g_Fusobacterium, g__Acinetobacter,* and *g_Bacteroides* were more abundant in colostrum samples of multiparous cows than in primiparous cows. The prevalence of *Staphylococcus* spp. during the periparturient period in primiparous and multiparous cows was investigated by Matthews (1992) and, in agreement with our findings, they found *Staphylococcus* spp. to be more prevalent in primiparous than in multiparous samples (Matthews et al., 1992). A recently published study described differences in PCR-based bacteriological results between samples

obtained directly from the udder cistern using a needle and vacuum tube, and samples obtained conventionally (Hiitiö et al., 2016). In the conventional samples, the number of positive *Staphylococcus* spp. was more than twice that of samples taken with the needle technique, indicating that most of the *Staphylococcus* spp. could have originated from the teat skin or environmental sources. Although in the present study we performed the disinfection procedure correctly, contamination from the teat skin could still have occurred and consequently affected the microbial taxonomic structure described herein. However, the presence of strict anaerobic bacteria such as *g_Fusobacterium*, *g_Prevotella*, *g_Bacteroides,* and *o_Clostridiales* in our study indicates that the complex colostrum microbial community that we described is likely an accurate representation of the colostrum microbiome, and not merely a result of teat skin contamination.

Our results showed that CMC samples from primiparous cows were less diverse than NCMC samples from cows of this same parity. However, the same was not the case for multiparous animals. Reduced microbial diversity has been linked to several diseases in humans, such as obesity (Turnbaugh and Gordon et al., 2009) and bowel disease (Qin et al., 2010); however higher microbial diversity has been detected in vaginosis (Fredricks et al., 2005) and respiratory tract illnesses (Charlson et al., 2011). Thus, different niches of the body appear to have different bacterial diversity profiles. In our study, the colostrum sampled from primiparous quarters that eventually developed CM was less diverse than the colostrum from quarters that did not experience mastitis during the first 30 days postpartum, suggesting that, in primiparous colostrum, lower diversity may be associated with future disease events.

At the phylum level, Tenericutes and Fusobacterium were more abundant in CMC than in NCMC of primiparous animals. Both of these phyla contain important bacterial species, such as

Mycoplasma spp. (Fox et al., 2005) and *Fusobacterium necrophorum* (Tan et al., 1996, Oikonomou et al., 2012), which have been associated with the etiopathogenesis of CM. In agreement with the phylum-level results, *Mycoplasma* and *Fusobacterium* genera were more abundant in CMC than in NCMC of primiparous cows. Thus, our data suggest that the colostrum microbiome of primiparous animals which eventually developed early CM harbors lower bacterial diversity, and that there are taxonomic changes at the phylum and genus levels which are associated with mammary gland health; the same was not observed for multiparous cows.

The core microbiome has been described as the population of microbes that is conserved regardless of host genetics and diet, and its disruption could indicate metabolic imbalances and disease of the gut (Ley et al., 2006, Turnbaugh and Gordon et al., 2009). In our study, the core microbiome was defined as bacterial taxa shared by all colostrum samples of all cows examined, regardless of parity or colostrum category (CMC or NCMC). Twenty taxa were detected as being part of the colostrum core microbiome, including *g__Staphylococcus*, *g__Mycoplasma*, and *g__Corynebacterium*. Some members of these bacterial genera such as *Staphylococcus chromogenes* (Pyörälä and Taponen et al., 2009) and *Staphylococcus aureus* (Fox et al., 2001), *Mycoplasma bovis* (Fox and Gay et al., 1993, Fox et al., 2005, Fox et al., 2012), and *Corynebacterium bovis* (Fox and Gay et al., 1993) have been systematically associated with mastitis. Thus, further studies are needed to better understand the complexity of the colostrum core microbiome and its classification as a taxonomic marker of mammary gland health.

CONCLUSION

In conclusion, we showed that the bovine colostrum contains a highly diverse and rich microbial community. Primiparous cow colostrum was shown to harbor a richer microbiota than

that of multiparous cows, and differences in the microbial community structure between parity were also observed. Differences in diversity and microbial community structure between NCMC and CMC existed for primiparous cows; however, the same was not detected for multiparous cows.

ACKNOWLEDGEMENTS

The present research project was supported by the Agriculture and Food Research Initiative Competitive Grant no. 2013-67015-21233 from the USDA National Institute of Food and Agriculture (Washington, DC.).

REFERENCES

Aarestrup, F. M. and N. E. Jensen. 1997. Prevalence and duration of intramammary infection in Danish heifers during the peripartum period. J Dairy Sci 80(2):307-312.

Addis, M. F., A. Tanca, S. Uzzau, G. Oikonomou, R. C. Bicalho, and P. Moroni. 2016. The bovine milk microbiota: insights and perspectives from -omics studies. Mol Biosyst.

Amann, R. I., W. Ludwig, and K. H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59(1):143-169.

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., Mills, D. A., Caporaso, J. G. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Methods 10(1):57-59.

Bradley, A. J. and M. J. Green. 2000. A study of the incidence and significance of intramammary enterobacterial infections acquired during the dry period. J Dairy Sci 83(9):1957-1965.

Bradley, A. J. and M. J. Green. 2004. The importance of the nonlactating period in the epidemiology of intramammary infection and strategies for prevention. Vet Clin North Am Food Anim Pract 20(3):547-568.

Cabrera-Rubio, R., M. C. Collado, K. Laitinen, S. Salminen, E. Isolauri, and A. Mira. 2012. The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery. Am J Clin Nutr 96(3):544-551.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., Knight, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8):1621-1624.

Capuco, A. V., R. M. Akers, and J. J. Smith. 1997. Mammary growth in Holstein cows during the dry period: quantification of nucleic acids and histology. J Dairy Sci 80(3):477-487.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. Scand J Stat. 11:265-270.

Charlson, E. S., K. Bittinger, A. R. Haas, A. S. Fitzgerald, I. Frank, A. Yadav, F. D. Bushman, and R. G. Collman. 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med 184(8):957-963.

Chase, C. C., D. J. Hurley, and Reber, A. J. 2008. Neonatal immune development in the calf and its impact on vaccine response. Vet Clin North Am Food Anim Pract 24(1):87-104.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460-2461.

Elfstrand, L., Lindmark-Månsson, H., Paulsson, M., L. Nyberg, and Åkesson, B. 2002. Immunoglobulins, growth factors and growth hormone in bovine colostrum and the effects of processing. Int Dairy J 12(11):879-887.

Falentin, H., L. Rault, A. Nicolas, D. S. Bouchard, J. Lassalas, P. Lamberton, J. M. Aubry, P. G. Marnet, Y. Le Loir, and S. Even. 2016. Bovine Teat Microbiome Analysis Revealed Reduced Alpha Diversity and Significant Changes in Taxonomic Profiles in Quarters with a History of Mastitis. Front Microbiol 7:480.

Federation, I. D. 1987. Machine Milking and Mastitis. Bulletin of the International Dairy Federation.

Fox, L. K. 2012. *Mycoplasma* mastitis: causes, transmission, and control. Vet Clin North Am Food Anim Pract 28(2):225-237.

Fox, L. K., J. H. Kirk, and A. Britten. 2005. *Mycoplasma* mastitis: a review of transmission and control. J Vet Med B Infect Dis Vet Public Health 52(4):153-160.

Fox, L. K., Bayles, K. W., & Bohach, G. A. 2001. *Staphylococcus aureus* mastitis. Pages 271-294 in *Staphylococcus aureus* Infection and Disease. Springer US.

Fox, L. K. and J. M. Gay. 1993. Contagious mastitis. Vet Clin North Am Food Anim Pract 9(3):475-487.

Fredricks, D. N., T. L. Fiedler, and J. M. Marrazzo. 2005. Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353(18):1899-1911.

Gilbert, J. A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C. T., Desai, N., Eisen, J. A., Evers, D., Field, D., W. Feng, Huson, D., Jansson, J., Knight, R., Knight, J., Kolker, E., Konstantindis, K., Kostka, J., Kyrpides, N., and R. Mackelprang, McHardy, A., Quince, C., Raes, J., Sczyrba, A. Shade, and A., Stevens, R. 2010. Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. Stand Genomic Sci 3(3):243-248.

Green, M. J., L. E. Green, G. F. Medley, Y. H. Schukken, and A. J. Bradley. 2002. Influence of dry period bacterial intramammary infection on clinical mastitis in dairy cows. J Dairy Sci 85(10):2589-2599.

Hammon, H. M., J. Steinhoff-Wagner, J. Flor, U. Schönhusen, and C. C. Metges. 2013. Lactation Biology Symposium: role of colostrum and colostrum components on glucose metabolism in neonatal calves. J Anim Sci 91(2):685-695.

Hiitiö, H., H. Simojoki, P. Kalmus, J. Holopainen, S. Pyörälä, and S. Taponen. 2016. The effect of sampling technique on the PCR-based bacteriological results of bovine milk samples. J Dairy Sci.

Khodayar-Pardo, P., L. Mira-Pascual, M. C. Collado, and C. Martínez-Costa. 2014. Impact of lactation stage, gestational age and mode of delivery on breast milk microbiota. J Perinatol 34(8):599-605.

Kuehn, J. S., P. J. Gorden, D. Munro, R. Rong, Q. Dong, P. J. Plummer, C. Wang, and G. J. Phillips. 2013. Bacterial community profiling of milk samples as a means to understand culture-negative bovine clinical mastitis. PLoS One 8(4):e61959.

Ley, R. E., P. J. Turnbaugh, S. Klein, and J. I. Gordon. 2006. Microbial ecology: human gut microbes associated with obesity. Nature 444(7122):1022-1023.

Lindner, J. D. D., Santarelli, M., Yamaguishi, C. T., Soccol, C. R., and Neviani, E. 2011. Recovery and identification of bovine colostrum microflora using traditional and molecular approaches. Food Technol Biotech 49(3):364.

Matthews, K. R., Harmon, R. J. and B. E. Langlois. 1992. Prevalence of *Staphylococcus* species during the periparturient period in primiparous and multiparous cows. J Dairy Sci 75(7):1835-1839.

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., Andersen, G. L., Knight, R., and P. Hugenholtz. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610-618.

Natzke, R. P. 1981. Elements of mastitis control. J Dairy Sci 64(6):1431-1442.

NMC. 1999. Laboratory Handbook on Bovine Mastitis: National Mastitis Council. Page 222.

NRC. 2001. Nutrient Requirements of Dairy Cattle. Natl. Acad. Press, Washington, DC.

Ogra, S. S., D. Weintraub, and P. L. Ogra. 1977. Immunologic aspects of human colostrum and milk. III. Fate and absorption of cellular and soluble components in the gastrointestinal tract of the newborn. J Immunol 119(1):245-248.

Oikonomou, G., M. L. Bicalho, E. Meira, R. E. Rossi, C. Foditsch, V. S. Machado, A. G. Teixeira, C. Santisteban, Y. H. Schukken, and R. C. Bicalho. 2014. Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. PLoS One 9(1):e85904.

Oikonomou, G., V. S. Machado, C. Santisteban, Y. H. Schukken, and R. C. Bicalho. 2012. Microbial diversity of bovine mastitic milk as described by pyrosequencing of metagenomic 16s rDNA. PLoS One 7(10):e47671.

Oviedo-Boyso, J., J. J. Valdez-Alarcón, M. Cajero-Juárez, A. Ochoa-Zarzosa, J. E. López-Meza, A. Bravo-Patiño, and V. M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J Infect 54(4):399-409.

Pakkanen, R., and J. Aalto. 1997. Growth factors and antimicrobial factors of bovine colostrum. Int Dairy J 7(5):285-297.

Pyörälä, S. and S. Taponen. 2009. Coagulase-negative staphylococci-emerging mastitis pathogens. Vet Microbiol 134(1-2):3-8.

Qin, J., R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J. M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, M. Jian, Y. Zhou, Y. Li, X. Zhang, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, P. Bork, S. D.

Ehrlich, and M. Consortium. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464(7285):59-65.

Rasolofo, E. A., D. St-Gelais, G. LaPointe, and D. Roy. 2010. Molecular analysis of bacterial population structure and dynamics during cold storage of untreated and treated milk. Int J Food Microbiol 138(1-2):108-118.

Rollin, E., K. C. Dhuyvetter, and M. W. Overton. 2015. The cost of clinical mastitis in the first 30 days of lactation: An economic modeling tool. Prev Vet Med 122(3):257-264.

Schloss, P. D. and J. Handelsman. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71(3):1501-1506.

Tan, Z. L., T. G. Nagaraja, and M. M. Chengappa. 1996. *Fusobacterium necrophorum* infections: virulence factors, pathogenic mechanism and control measures. Vet Res Commun 20(2):113-140.

Taponen, S., L. Salmikivi, H. Simojoki, M. T. Koskinen, and S. Pyörälä. 2009. Real-time polymerase chain reaction-based identification of bacteria in milk samples from bovine clinical mastitis with no growth in conventional culturing. J Dairy Sci 92(6):2610-2617.

Turnbaugh, P. J. and J. I. Gordon. 2009. The core gut microbiome, energy balance and obesity. J Physiol 587(Pt 17):4153-4158.

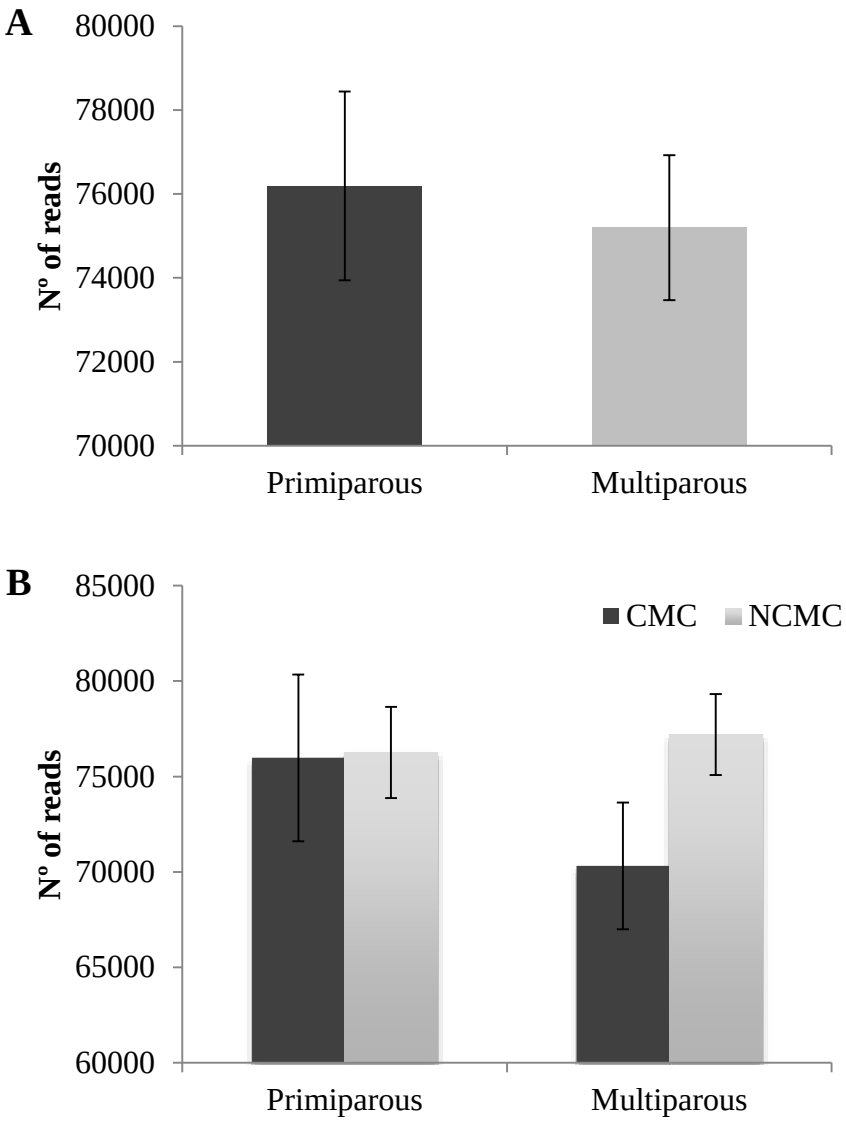
USDA. 2007. Dairy 2007: Biosecurity Practices on U.S. Dairy Operations. USDA, Animal and Plant Health Inspection Service, Veterinary Services, Center for Epidemiology and Animal Health, Fort Collins, CO.

Ward, T. L., S. Hosid, I. Ioshikhes, and I. Altosaar. 2013. Human milk metagenome: a functional capacity analysis. BMC Microbiol 13:116.

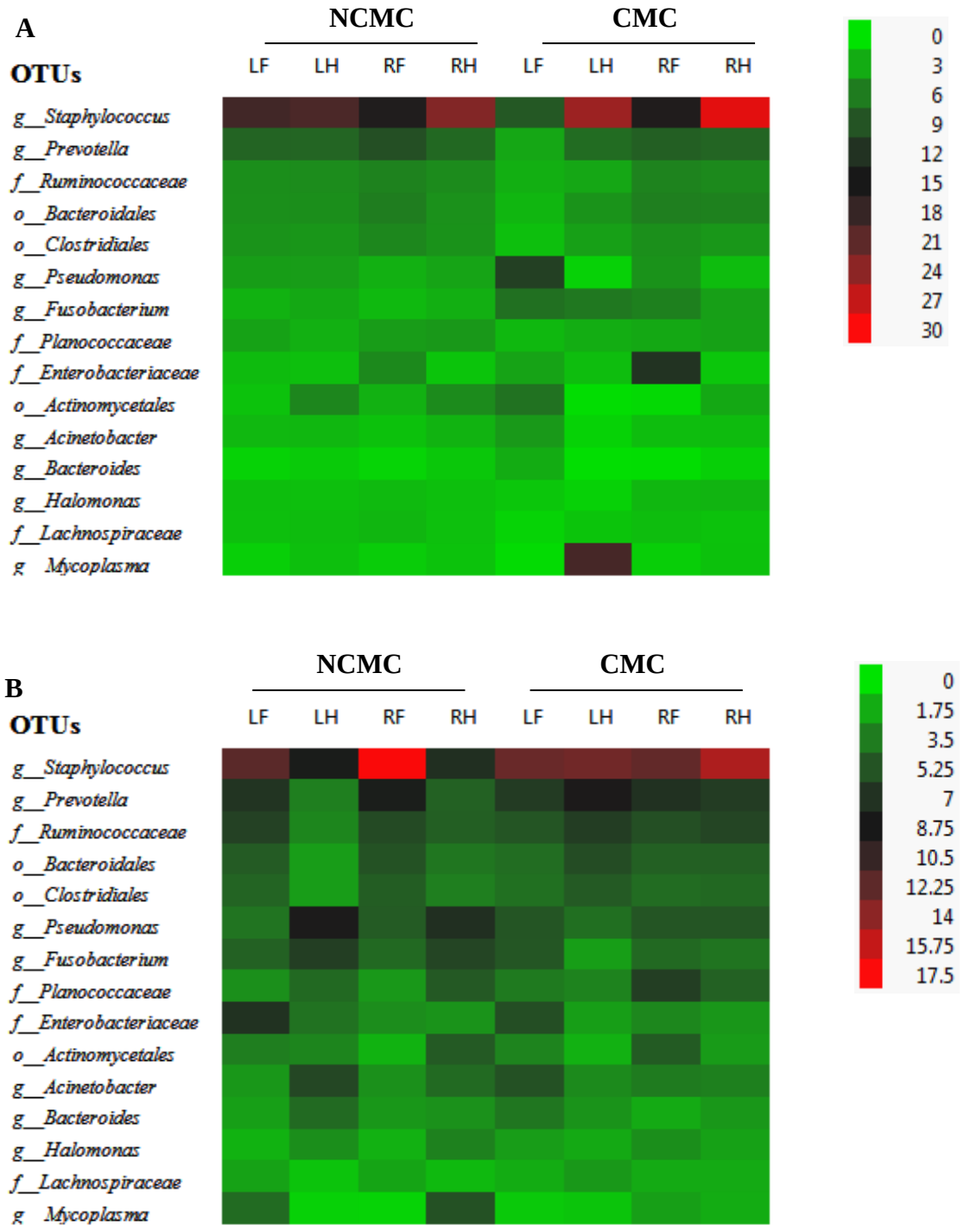
Watters, R. D., J. N. Guenther, A. E. Brickner, R. R. Rastani, P. M. Crump, P. W. Clark, and R. R. Grummer. 2008. Effects of dry period length on milk production and health of dairy cattle. J Dairy Sci 91(7):2595-2603.

Zand, S.M., 1976. Indexes associated with information theory in water quality. J Water Pollut Control Fed. 2026-2031.

Supplemental 3.1: Number of reads according to parity (A), and colostrum category (B). Error bars are positioned around the means and represent the standard error of the mean. † represents $0.05 \leq P < 0.1$ and * represent $P < 0.05$.



Supplemental 3.2: Heatmap illustrating the 15 most common microbial taxa detected in colostrum samples of primiparous (A) and multiparous cows (B) according to cow udder quarter and colostrum category (non-clinical mastitic colostrum & clinical mastitic colostrum). *o* (order), *f* (family), and *g* (genus) represent the taxonomic level of the described bacterium. Each square in the heatmap represents the abundance level of a single category. Small relative abundance values are light green, progressing to higher values as black and red.



CHAPTER FOUR

THE UPPER RESPIRATORY TRACT MICROBIOME AND ITS POTENTIAL ROLE IN BOVINE RESPIRATORY DISEASE AND OTITIS MEDIA

Svetlana F. Lima[1], Andre Gustavo V. Teixeira[1], Catherine H. Higgins[1], Fabio S. Lima[2], Rodrigo C. Bicalho[1]*

Nature Scientific Reports. *6: 1-6*

[1]Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

[2]Department of Veterinary Clinical Medicine, College of Veterinary Medicine, University of Illinois, Urbana, Illinois, USA

ABSTRACT

The upper respiratory tract (URT) hosts a complex microbial community of commensal microorganisms and potential pathogens. Analyzing the composition and nature of the healthy URT microbiota and how it changes over time will contribute to a better understanding of the pathogenesis of pneumonia and otitis. A longitudinal study was conducted including 174 Holstein calves that were divided in four groups: healthy calves, calves diagnosed with pneumonia, otitis or both diseases. Deep pharyngeal swabs were collected on days 3, 14, 28, and 35 of life, and next-generation sequencing of the 16S rRNA gene as well as quantitative PCR was performed. The URT of Holstein dairy calves aged 3 to 35 days revealed to host a highly diverse bacterial community. The relative abundances of the bacterial genera *Mannheimia*, *Moraxella*, and *Mycoplasma* were significantly higher in diseased versus healthy animals, and the total bacterial load of newborn calves at day 3 was higher for animals that developed pneumonia than for healthy animals. Our results corroborate the existing knowledge that species of *Mannheimia* and *Mycoplasma* are important pathogens in pneumonia and otitis. Furthermore, they suggest that species of Moraxella can potentially cause the same disorders (pneumonia and otitis), and that high neonatal bacterial load is a key contributor to the development of pneumonia.

INTRODUCTION

Bovine respiratory disease (BRD) is a complex, multifactorial disorder caused by a combination of microbial pathogens (Angen et al., 2009, Confer et al., 2009, Pardon et al., 2011), impaired host immunity (Lago et al., 2006, Chase et al., 2008, Gorden and Plummer et al., 2010), environmental factors (Webster et al., 1983, Lago et al., 2006, Gorden and Plummer et al., 2010), and inadequate housing conditions (Lago et al., 2006, Gorden and Plummer et al., 2010). Despite advances in veterinary medicine and technology to control BRD, it remains a huge economic burden for both the dairy and beef industries due to calf mortality, treatment expenses, and additional labor incurred. Furthermore, BRD has substantial long-term consequences on performance by negatively impacting growth (Virtala et al., 1996), reproductive performance (Waltner-Toews, 1986), and longevity (Waltner-Toews et al., 1986, Warnick et al., 1995). In the USA dairy industry, BRD is a major contributor to mortality and morbidity(USDA, 2007). The United States Department of Agriculture National Animal Health Monitoring Service (NAHMS, 2007) reported that BRD affects 12.4% of calves during the pre-weaning period and is responsible for 22.5% of the mortality documented during the same period (USDA, 2007). Additionally, 5.9% of post-weaning animals are eventually diagnosed with BRD, which is responsible for 46.5% of the mortality documented during the same period. The detrimental economic impact of BRD on the American beef industry is even larger than on the dairy industry. BRD is considered to be the most expensive disease affecting feedlot cattle and it has been estimated to cause losses of approximately one billion dollars per year in the USA (Fulton, 2002, Griffin et al., 2010).

Several viral and bacterial etiological agents have been associated with bovine respiratory tract (RT) disease. Bovine viral diarrhea virus (BVDV), bovine respiratory syncytial virus

(BRSV), and parainfluenza type 3 virus (PI-3) have all been described as important causative agents of the BRD complex (Chase et al., 2008, Griffin et al., 2010). It is thought that primary viral infection may render the RT epithelium more susceptible to bacterial colonization. Viral infections can impair the epithelial layer of the RT mucosa by disarranging host cellular functions and/or killing infected epithelial cells, thus exposing the RT basement membrane. Furthermore, viruses may damage ciliated host cells, resulting in reduced mucociliary velocity (reduction of bacterial clearance) and thus leading to a compromised host immune response to secondary bacterial infection (Babiuk et al., 1988, Czuprynski et al., 2004, Bosch et al., 2013). The key bacterial pathogens associated with bovine pneumonia include *Pasteurella multocida* (Thomson et al., 1975, Allen, 1991), *Mannheimia haemolytica* (Thomson et al., 1975, Haines et al., 2001), *Histophilus somni* (Shahriar et al., 2002), *Mycoplasma bovis* (Haines et al., 2001, Shahriar et al., 2002), and other *Mycoplasma* spp. (Pardon et al., 2011).

Another relatively common disease affecting dairy calves is otitis media. This infection of the middle ear manifests as head tilt and sometimes facial paralysis due to involvement of cranial nerves VII and VIII and peripheral vestibular structures (Jensen et al., 1983, Yeruham et al., 1999, Van Biervliet et al., 2004). A recent study evaluated the impact of BRD, diarrhea, arthritis, and otitis on mortality and carcass traits in white veal calves, and an increased mortality risk was found for otitis (Hazard ratio equal to 7.0) (Pardon et al., 2013). It has been suggested that otitis media and pneumonia may evolve from URT infection, because anatomically the nasopharynx area communicates with the nose cavity, the sinuses, the middle ears, and the larynx, and URT-resident microbes can be a source for lower respiratory tract infections (García-Rodríguez and Fresnadillo Martínez et al., 2002, Murphy et al., 2009). This is also true for middle ear infections, as the nasopharynx is connected to the middle ear via the Eustachian tube

(Murphy et al., 2009), thereby supporting a strong link between the URT and both otitis and pneumonia. Moreover, both diseases affect calves of the same age and share common risk factors, and the most dominant bacteria reportedly involved in the etiology of bronchopneumonia are also associated with otitis media (Jensen et al., 1983). Pathogens commonly associated with otitis media include *Mannheimia haemolytica* (Yeruham et al., 1999), *Histophilus somni* (McEwen and Hulland, 1985), *Mycoplasma* spp. (Francoz et al., 2004), *Mycoplasma bovis* (Walz, 1997, Maeda, 2003), *Pasteurella multocida* (Jensen et al., 1983, Baba et al., 1988), *Staphylococcus* spp. (Yeruham et al., 1999, Arcangioli et al., 2012), and *Streptococcus* spp. (Baba et al., 1988).

Currently, intensive efforts have focused on understanding the composition and nature of bodily microbial populations in a balanced microbiome state, and how shifts in such microbial community structures impact the health of both humans (Angen et al., 2009, D'Argenio and Salvatore et al., 2015) and animals (Khafipour et al., 2009, Oikonomou et al., 2012). According to Bosch et al. (Bosch et al., 2013), imbalances of the upper respiratory tract (URT) ecosystem may result in invasion by and overgrowth of bacterial pathogens, leading to respiratory disease. A recent study by Homan et al. (Holman et al., 2015) assessed the URT microbiome of feedlot cattle on the day of arrival and again 60 days later and found significant differences between the URT microbiomes at the two time points. Unfortunately, comparisons between healthy and sick animals were not evaluated in that study (Holman et al., 2015). To our knowledge, the URT microbiome of dairy calves has not been previously investigated using modern microbiome techniques. The 16S rRNA gene library-based molecular strategy is a powerful approach for identifying members of a microbial community and quantifying their relative abundance while avoiding the limitations imposed by culture-dependent methods and biochemical approaches

(Amann et al., 1995, Schloss and Handelsman et al., 2005). Furthermore, it is a rapid and cost-effective method for assessing bacterial diversity and is a useful tool for pathogen discovery and identification (Kolbert and Persing et al., 1999).

Therefore, the goal of this study was to characterize longitudinally the URT microbiome of Holstein dairy calves by using high-throughput sequencing of the 16S rRNA gene. We aimed to compare the URT microbial communities of healthy and unhealthy subjects at each time point of our data collection.

MATERIALS AND METHODS

Ethics statement

This study was conducted on a commercial dairy farm located near Ithaca, New York, from November 2013 until February of 2014. The collection of nasal swabs samples from calves at days 3, 14, 28 and 35 of life was authorized by the farm owner. Animal Care and Use Procedures were cared for according to the guidelines set by Dairy Cattle Husbandry (nº 518) [72]. "The Animal Care and Use Procedures are produced and enforced to ensure the welfare of animals used in research and teaching at Cornell University". All experimental protocols using cattle were reviewed and approved by the Institutional Animal Care and use Committee of Cornell University (Protocol number: 2013-0076).

Animals and facilities

Pregnant cows at stage 1 or 2 of parturition were transferred from the close-up free-stall barn into two maternity pens (400 m^2 deep-bedded pens). After parturition, calves were removed from the maternity pen and placed into a newborn pen bedded with dry sawdust and heated with

heating lamps during the winter months. Colostrum from multiparous and primiparous cows was pooled and used in the study. All calves were fed approximately 4 L of raw colostrum at once by an esophageal feeder (Oral Calf Feeder Bag with Probe, Jorvet) within 4 hours of birth.

Twice daily, newborn calves were allocated from the newborn pen to the calf barn. The calf barn was a greenhouse type of barn with positive ventilation and divided into 18 identical group-pens. Group-pens had a total area of 70 m^2 and were bedded with straw bedding on top of a thin layer of dry composted manure. Steel gates divided the group pens, and calves were allocated by birth order into each pen until the pen was completely full (a total of 25 calves per pen). All calves remained in the same pen from day 1 of life until fully weaned (approximately 65 days). Birth weight and weight at weaning of all heifer calves were measured by farm employees using a Waypig 15, 62-inch digital scale (Waypig-15, Vittetoe Inc., Keota, IA).

Calves were fed ad-libitum acidified non-saleable milk. The feeding system was fully automated. Briefly, the acidification was performed inside a sealed stainless-steel tank where the non-saleable cold (5^{o}C) milk was mixed continuously with organic acid until pH 4.5 was reached. The acidified milk was kept for 72 hours inside the stainless-steel tank after the acidification process was finished. Then, the milk was directed to a smaller stainless-steel tank, which maintained the milk at a warm temperature and distributed it to the pens. To support the ad-libitum system, 6 nipples per pen were connected to the smaller tank and the acidified non-saleable milk was available from day 1 to day 55 of life, when a reduction of milk availability was initiated. All calves in this study were weaned by reducing the milk availability starting on day 55 until complete absence of acidified non-saleable milk at 65 days of life.

Deep pharyngeal swab collection

A cohort of 174 Holstein heifer calves was selected randomly for this study. Deep pharyngeal swabs were performed on days 3, 14, 28, and 35 of life using a 20-cm DNA-free sterile swab (Puritan Medical Products, Guilford, ME) covered by a thin sterile plastic sheath. Prior to sampling, the selected calves were appropriately restrained and the nostril was cleaned using a paper towel. The plastic-covered sterile swab was inserted into the right nasal cavity at a depth of approximately 15 cm, the plastic sheet was then broken, exposing the swab to the URT mucosa, and a 360^{o} rotation was performed to better standardize sample collection. The swab was then retracted back into the plastic sheath and removed. The tip of the swab was placed inside a sterile plastic tube and labeled. Samples were kept on ice until they were transferred to the laboratory at Cornell Veterinary School and stored at -20 °C until further processing.

Case definition

Pneumonia was defined when two or more of the following clinical signs were detected in a calf: cough, rectal temperature >39.5 °C, respiratory rate >40 breaths/min, increased cranioventral lung sounds or wheezes. Otitis was defined by observation of ear pain evidenced by head shaking, scratching or rubbing the ears, epiphora, ear droop, signs of facial nerve paralysis, with or without fever (rectal temperature >39.5 °C).

Two dedicated farm employees with over 10 years of experience and trained by Cornell University veterinarians (Ambulatory and Production Medicine Department), were responsible for overseeing the calf facility and making the initial detection of pneumonia and otitis. When farm employees detected animals that were displaying signs of disease such as depression, inappetence, dehydration, increased respiratory rate, or a head tilt (otitis) a full physical examination was carried out to determine the diagnosis of the disease. Once the farm employees

examined the affected calves an experienced veterinarian member of the research team performed a second confirmatory physical examination. Calves diagnosed with pneumonia and/or otitis were treated according to standard farm protocol (Resflor Gold, Merck Animal Health).

DNA extraction

Isolation of DNA from swabs of the URT was performed by adding 1.5 ml of DNA-free water into a 2-ml microcentrifuge tube containing a nasal swab sample, which was disrupted using a Mini-Beadbeater-8 (Biospec Products, Battersville, OK). Swabs were removed from the microcentrifuge tubes and the remaining liquid was centrifuged for 10 minutes at 13,000 rpm. The supernatant was discarded and the DNA was extracted from the pellet using the PowerSoil DNA Isolation Kit (MO BIO Laboratory Inc., Carlsbad, CA). DNA concentration and purity were evaluated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 230, 260, and 280 nm.

Quantitative PCR

In order to determine the total bacterial load of the URT samples, we cloned a plasmid containing the amplified V6 hypervariable region into TOP10 cells by using a Zero Blunt® TOPO® PCR cloning kit (Life Technologies, Darmstadt, Germany). Plasmid was purified with a QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA) and quantified using Quant-iT™ PicoGreen® and a dsDNA Broad Range Assay Kit (Life Technologies Corporation, Carlsbad, CA). Insertion was confirmed by agarose gel electrophoresis, and by sequencing the plasmid at the Cornell University Life Science Core Laboratories Center.

16S rRNA gene copy numbers were measured by quantitative PCR (qPCR) using Unibac primer (forward: 50-TGG AGC ATG TGG TTT AAT TCG A-30; reverse: 50-TGC GGG ACT TAA CCC AAC A-3) (Nonnenmacher et al., 2004, Boutin, 2015). PCRs were performed in 15 µL volumes composed of 1X iQ^{TM}Sybr Green Supermix (BIO-RAD Laboratories, Hercules, CA), 300nM of each primer and 50ng-5pg of genomic DNA (or plasmid DNA standards). The thermal cycler conditions were as follows: denaturation at 95°C for 3 min, 40 amplification cycles (95°C for 10 s, 55°C for 30 s) and two final steps at 95°C for 1 min and 55°C for 1 min followed by melting curve determination. All reactions were performed in duplicate using a MyiQ^{TM} Real-Time PCR Detection System (BIO-RAD Laboratories, Hercules, CA). Quantification of 16S rRNA target DNA was achieved by 10-fold serial dilutions ranging from 10^0 to 10^7 plasmid copies of the previously quantified plasmid standards. Plasmid standards and URT samples were run in duplicates. The average of the cycle threshold value was used for calculation of the total bacterial load.

PCR amplification of the V4 hypervariable region of bacterial 16S rRNA genes

The 16S rRNA gene was amplified by PCR from individual metagenomic DNA samples from the URT using barcoded primers. For amplification of the V4 hypervariable region of the bacterial/archaeal 16S rRNA gene, primers 515F and 806R were used according to previously described methods and optimized for the Illumina MiSeq platform (Caporaso, 2012). The Earth Microbiome Project (http://www.earthmicrobiome.org/) (Gilbert et al., 2010) was used to select 140 different 12-bp error-correcting Golay barcodes for the 16S rRNA gene PCR, as previously described (Caporaso, 2012). The 5'-barcoded amplicons were generated in triplicate using 12-300 ng of template DNA, 1X EconoTaq® Plus Green Master Mix (Lucigen®, Middleton, WI) and 10 µM of each primer. The PCR conditions for amplification of the 16S rRNA gene included an

initial denaturing step of 94°C for 3 min followed by 35 cycles of 94°C for 45 s, 50°C for 1 min and 72°C for 90 s and a final elongation step of 72°C for 10 min. Before sequencing, replicate amplicons were pooled and purified with a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) and visualized by electrophoresis using 1.2% (wt/vol) agarose gels stained with 0.5 mg/ml ethidium bromide. Blank controls in which no DNA was added to the reaction were performed. Purified amplicon DNA was quantified using Quant-iT™ PicoGreen® and a dsDNA Broad Range Assay Kit (Life Technologies Corporation, Carlsbad, CA).

Sequencing, bioinformatics, and statistical analysis

Aliquots of URT amplicon samples were standardized to the same concentration and pooled into 5 different sequencing runs according to individual barcode primers for the 16S rRNA gene. Final equimolar libraries were sequenced using the MiSeq reagent kit v2 (300 cycles) on the MiSeq platform (Illumina, Inc., San Diego, CA). The generated 16S rRNA gene sequences were processed through the open source software pipeline Quantitative Insights Into Microbial Ecology (QIIME) version 1.7.0-dev (Caporaso, 2012). Sequences were filtered for quality using established guidelines (Bokulich, 2013). Sequences were binned into OTUs based on 97% identity using UCLUST (Edgar, 2010) against the Greengenes reference database (McDonald, 2012), May 2013 release. Low-abundance clusters were filtered and chimeric sequences were removed using USEARCH (Edgar, 2010). The representative sequences for each OTU were compared against the Greengenes database for taxonomy assignment, and only full-length, high-quality reads (-r = 0) were used for analysis. Additionally, we generated a species-level OTU table using the MiSeq Reporter Metagenomics Workflow. The MiSeq Reporter classification is based on the Greengenes database (http://greengenes.lbl.gov/) and the output of

this workflow is a classification of reads at multiple taxonomic levels: kingdom, phylum, class, order, family, genus, and species.

Shannon and Chao1 indexes output were generated by QIIME pipeline. Before estimating the Shannon and Chao1 indexes, all sample libraries were rarefied to an equal depth of 10,000 sequences using QIIME. Chao1 and Shannon indexes, total number of reads, and the log of the 16S rDNA copy number (total bacterial load) were analyzed using repeated measures ANOVA by general linear models fitted in JMP Pro 11 (SAS Institute Inc., Cary, NC). Dunnett's multiple comparisons procedure was performed to compare the mean number of reads, Shannon index and Chao 1 index of each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the healthy samples within each day of data collection (days 3, 14, 28, and 35).

Correlations between total bacterial load and alpha-diversity indexes (Shannon and Chao 1 indexes) were assessed using simple linear regression in JMP Pro 11 software (SAS Institute Inc.).

The relative abundances of microbial phyla and genera types in URT samples of calves at ages 3, 14, 28, and 35 within each health status were compared using general linear models fitted in JPM Pro 11 (SAS Institute Inc.). Dunnett's multiple comparisons procedure was used to compare the mean relative abundance of the most abundant bacterial phyla and the genera of each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the healthy samples within each day of data collection (days 3, 14, 28, and 35). Differences with a value of $P \leq 0.05$ were considered significant and those with a value of $0.05 < P \leq 0.10$ were considered tendencies.

Descriptive statistics for birth weight (BW) and average daily gain (ADG) were determined according to health status by using a general linear model (ANOVA) with JMP Pro

11 (SAS Institute Inc.). In total, 174 calves were enrolled in this study. Number of calves, disease incidence, mortality, BW and ADG (calculated by subtracting BW from the weaning weight and then dividing by days of life at weaning) during the pre-weaning period are presented in Table 1. Average age (days in life) at first diagnosis of pneumonia, otitis and pneumonia_otitis combined was assessed by using Distribution platform offered by JMP Pro 11 (SAS Institute Inc.).

RESULTS

Descriptive data

In total, 174 calves were enrolled in this study. Of these, 37 (21.3%) were diagnosed with pneumonia, 62 (35.6%) with otitis, 11 (6.3%) with pneumonia-otitis combined, and 64 (36.8%) were healthy (Table 1). The average age at first diagnosis was 22.5 days for pneumonia, 24.0 days for otitis, and 19.7 days for pneumonia-otitis combined (Table 4.1).

Table 4.1: Descriptive overview of the calves selected and enrolled in the study. Disease incidence, mortality, birth weight and average daily gain (ADG) are presented below. Calf birth weight and ADG are presented as means and standard error of the mean. Dunnett's multiple comparison procedure was used to compare ADG of each disease status (otitis, pneumonia and pneumonia-otitis combined) against the ADG of the "healthy" status.

	Healthy	Pneumonia	Otitis	Pneumonia and Otitis
n	64	37	62	11
Incidence (%)	36.8	21.3	35.6	6.3
Mortality (%)	1.15	0.57	1.72	0.57
Birth weight (kg)	38.05 (0.51)	38.36 (0.67)	38.52 (0.53)	36.16 (1.23)
ADG[1] (g/day)	659.4 (23.3)	633.9 (30.2)	577.4 (24.4)	564.6 (60.7)
ADL[2]	-	22.5 (1.3)	24 (0.9)	19.7 (2.0)

[1]ADG = Average daily gain was calculated by subtracting birth weight from the weaning weight and then dividing by days of life at weaning
[2]ADL = Average days of life at first diagnosis

Sequencing results

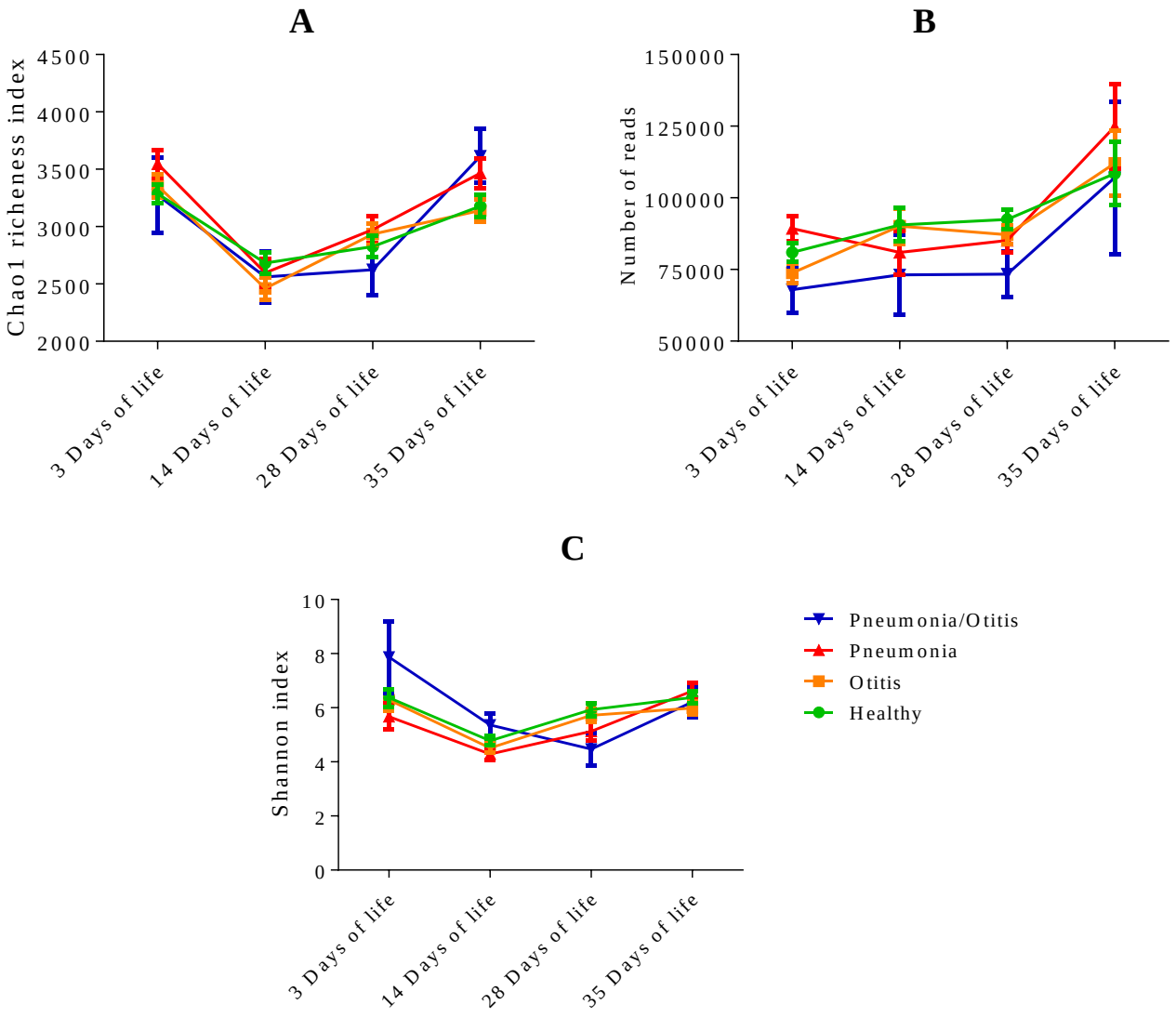
We collected 696 deep nasal swab samples from 174 Holstein calves. Samples were collected at days 3, 14, 28, and 35 of life. All samples collected were used individually to assess the microbiome by amplification and next-generation sequencing of the V4 region of the 16s rRNA gene. A total of 5 sequencing runs were performed using the Miseq sequencer (Illumina, Inc., San Diego, CA) and theV2 chemistry kits (300-cycles); approximately 140 barcoded samples were sequenced on each run.

Sequences were filtered for size, quality, and for the presence of chimeras and the total post-quality control number of sequences used in the study were 63,638,904. The average coverage was 91,567, the SD was 58,425, and the range was 1,423 to 657, 375 numbers of reads per sample.

Number of reads, richness and diversity indexes, and 16S rRNA gene copy numbers

The mean number of reads for each health status (healthy, pneumonia, otitis, pneumonia-otitis combined) was not significantly different within each time point (Figure 4.1A). Regarding OTU richness and diversity, the mean Chao1 richness index for each health status at each postnatal time point is illustrated in Figure 4.1 B. The mean Shannon diversity index for each health status at each postnatal time point is illustrated in Figure 4.1C. Chao1 richness index is a nonparametric estimator of the minimum richness and is based on the number of rare OTUs (singletons and doublets), within a sample (Chao, 1984). When a sample exhibits many singletons, it is likely that more undetected OTUs exist, and the Chao 1 richness index will estimate a higher richness than it would estimate for a sample deprived of rare OTUs. The Shannon diversity index accounts for both richness and abundance in a single value of evenness. Microbiomes that are numerically dominated by one or few organisms present low evenness, and when abundance is distributed equally among organisms the microbiome presents high evenness [42]. The richness and the evenness were analyzed to see whether any divergence is observed across health conditions. The Chao 1 and Shannon indexes did not differ significantly when comparing health statuses, regardless of the age time point (Figure 4.1B and C).

Figure 4.1: Bar graphs illustrating the Chao1 richness index (A), mean number of reads (B) and Shannon diversity index (C) for different postnatal ages. Error bars represent standard errors. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the status "healthy" within each sample collection date.



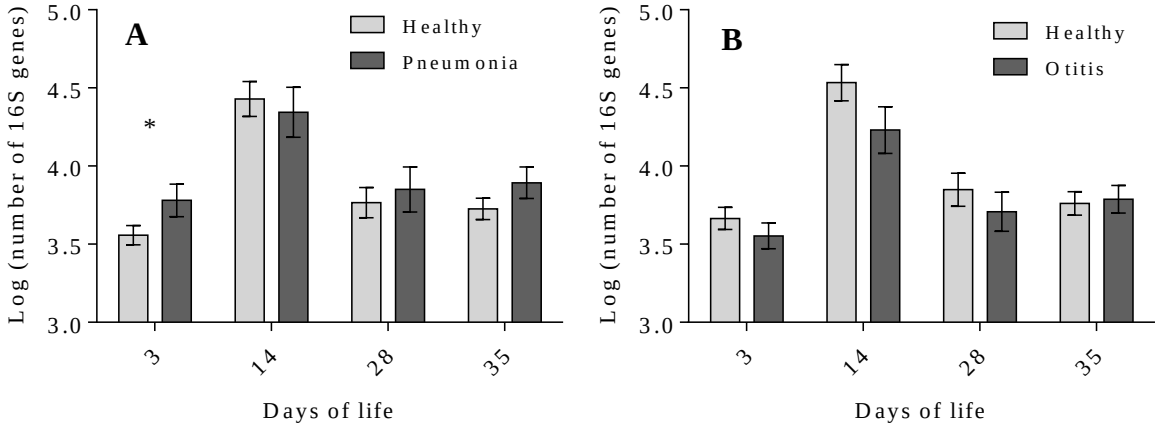
A negative correlation was detected between the total bacterial load, as assessed by the number of 16S rRNA gene copies, and the Shannon diversity index (r = -0.40, *P*-value<0.0001) (Supplemental 4.1). Similar negative correlations were found when data were stratified by the different disease statuses (Table 4.2). No correlation was found between the number of 16S rRNA gene copies and the Chao 1 richness index (r = -0.025, *P*-value = 0.20) (data not shown).

Table 4.2: Correlation between alpha-diversity and bacterial load in each heath status investigated. The bacterial load was measured via proxy of the number of 16S rRNA genes.

Health status	**Correlation**	***P* - values**
Healthy	-0.53	<0.0001
Pneumonia	-0.24	0.02
Otitis	-0.35	<0.001
Pneumonia-otitis combined	-0.48	0.01

Quantitative real-time PCR was used to monitor the amplification of the 16s rRNA targeted gene during PCR. As a result of this method an absolute quantification that gives the exact number of the target DNA molecules within a sample, by comparison with DNA standards (serial dilution of our 16S rRNA gene clone) using a calibration curve is provided. At day 3 of life, healthy calves had significantly lower total bacterial loads, as defined by the log10 copy numbers of the 16S rRNA gene, than calves diagnosed with pneumonia. The average counts of the 16s rRNA gene copies at day 3 was 3.80 $\log_{10}$ (SE = 0.10) for animals that were later diagnosed with pneumonia and 3.55 $\log_{10}$ (SE = 0.06) for healthy animals (Figure 4.2A).

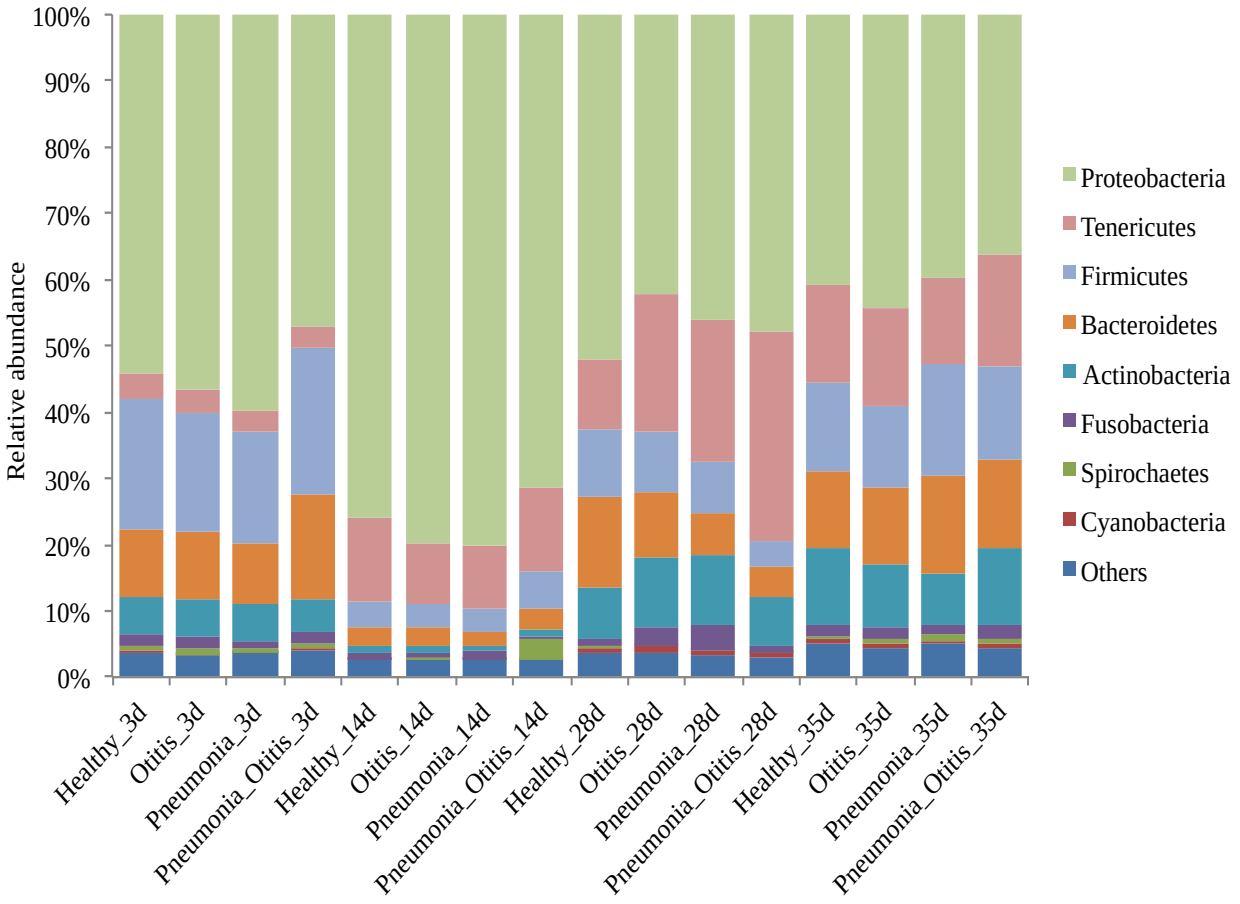
Figure 4.2: Mean log10 number of the 16S rRNA gene identified in upper respiratory tract samples of calves at various postnatal time points (3, 14, 28 and 35 days) and for different health statuses (healthy, pneumonia, and otitis). An asterisk between health statuses represents a significant difference ($P < 0.05$) for the age sampled.



Microbial phylum analysis

The relative abundances of the eight most common phyla of the URT regardless of age and health conditions (Proteobacteria, Tenericutes, Firmicutes, Bacteroidetes, Actinobacteria, Fusobacteria, Spirochaetes, and Cyanobacteria) are depicted in Figure 4.3. Proteobacteria was consistently the most abundant phylum across the four health categories (healthy, otitis, pneumonia, and pneumonia-otitis combined) (Figure 4.3).

Figure 4.3: Mean relative abundance of the most prevalent bacterial phyla identified in upper respiratory tract samples of calves at various postnatal time points (3, 14, 28 and 35 days) and for different health statuses (healthy, pneumonia, otitis, and pneumonia-otitis combined).



Bacterial genus analysis

The prevalence (% of animals detected with the respective OTU) of the 30 most abundant bacterial genera identified in the calf URT at days 3, 14, 24 and 35 of life is depicted in Table 4.3. The 30 most common bacterial genera and its respective mean relative abundance, identified throughout the different age time points and according to each health status evaluated are presented in Table 4. *Mannheimia*, *Mycoplasma*, *Moraxella*, *Psychrobacter*, and *Pseudomonas* were the top 5 genera regardless of health status (Table 4.4). Notably, *Mannheimia* and

Mycoplasma were the bacterial genera with the greatest increase in relative abundance over time (**Table 4.4 and Figure 4.4A and B**).

Table 4.3: Descriptive statistics of the 30 most abundant bacterial genera.

Genera	3 days		14 days		28 days		35 days	
	P[1]	MRA[2]	P	MRA	P	MRA	P	MRA
Mannheimia	100%	3.1	100%	24.9	100%	15.6	100%	15.4
Mycoplasma	100%	2.9	100%	10.8	100%	18.2	100%	15.5
Moraxella	100%	9.1	100%	18.4	99%	2.6	100%	4.8
Psychrobacter	100%	13.5	100%	8.4	99%	3.1	100%	7.0
Pseudomonas	100%	4.0	100%	13.1	100%	5.5	99%	1.2
Acinetobacter	100%	5.3	100%	1.7	100%	5.2	100%	3.6
Cellulomonas	98%	0.2	84%	0.1	98%	5.2	100%	5.2
Bacteroides	100%	3.3	99%	1.2	99%	2.3	100%	4.0
Escherichia	100%	4.2	99%	0.6	100%	2.9	99%	0.3
Corynebacterium	100%	2.5	99%	0.3	99%	2.1	100%	2.2
Fusobacterium	100%	1.1	100%	0.7	99%	1.9	100%	1.5
Pasteurella	100%	0.7	99%	2.2	99%	0.8	99%	0.6
Pedobacter	100%	1.1	99%	0.2	98%	2.3	100%	1.2
Streptococcus	100%	1.8	100%	1.5	99%	0.7	99%	0.5
Serratia	100%	2.6	99%	0.4	100%	1.4	99%	0.3
Prevotella	100%	1.0	99%	0.4	99%	0.8	100%	1.8
Staphylococcus	100%	1.1	99%	0.2	100%	1.8	100%	0.7
Ruminococcus	100%	1.3	99%	0.3	99%	0.7	100%	1.2
Candidatus Blochmannia	100%	0.9	100%	0.3	100%	1.1	100%	1.1
Porphyromonas	100%	1.3	99%	0.2	99%	0.4	100%	1.4
Clostridium	100%	1.9	100%	0.2	99%	0.4	100%	0.9
Blautia	100%	1.0	100%	0.2	99%	0.9	100%	1.2
Brenneria	100%	0.9	99%	0.8	96%	0.1	98%	0.5
Gallibacterium	100%	1.0	96%	0.5	99%	0.7	100%	0.4
Stenotrophomonas	99%	0.1	95%	0.1	99%	1.5	100%	0.4
Treponema	100%	0.9	100%	0.5	98%	0.2	99%	0.7
Sphingobacterium	100%	0.5	100%	0.5	99%	0.8	99%	0.4
Aggregatibacter	85%	0.1	99%	1.0	98%	0.1	97%	0.2
Aerococcus	99%	0.7	97%	0.2	99%	0.5	100%	0.4
Flavobacterium	100%	0.5	96%	0.1	99%	0.6	100%	0.5

[1]P = Percentage of study calves in which the indicated genus was detected at the given age at sample collection

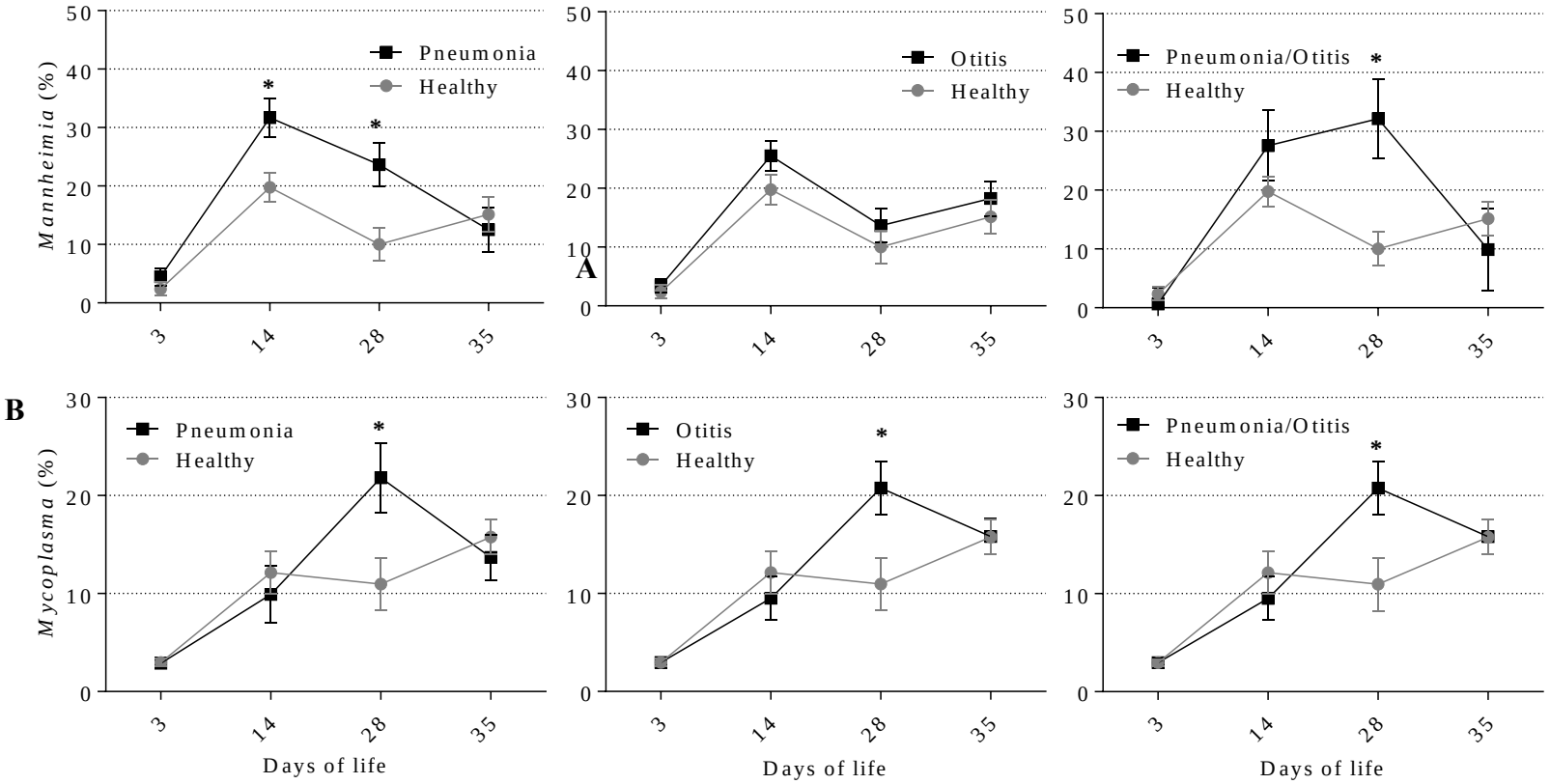
[2]MRA = Mean relative abundance of the respective bacterial genera at the given postnatal age

Table 4.4: The 20 most abundant genera detected in the URT and the correspondent abundance according to each heath conditions (healthy, otitis, pneumonia, pneumonia and otitis combined) and postnatal age (3, 14, 28 and 35 days of life). Dunnett's multiple comparison procedure was used to compare each disease status against the status "healthy" within each sample collection time point. [a,b,c] different superscripts among health status means statistical difference ($P < 0.05$) within each age sampled.

	3 days				14 days				28 days				35 days			
Genera	[1]H	[2]O	[3]P	[4]P&O	[1]H	[2]O	[3]P	[4]P&O	[1]H	[2]O	[3]P	[4]P&O	[1]H	[2]O	[3]P	[4]P&O
Mannheimia	2.4	3.4	4.5	0.7	19.8[a]	25.5[ab]	31.7[b]	27.6[ab]	10.0[a]	13.7[a]	23.7[b]	32.1[b]	15.2	18.3	12.5	9.9
Mycoplasma	3.0	3.0	2.9	3.0	12.1	9.5	9.9	13.3	11.0[a]	20.8[b]	21.8[b]	33.3[c]	15.8	15.8	13.7	17.7
Moraxella	8.0	11.0	8.9	5.1	13.5[a]	20.2[b]	21.9[b]	24.8[b]	1.9	2.1	4.3	4.3	3.2[a]	4.4	7.2[b]	8.4
Psychrobacter	12.8	14.4	13.5	11.9	7.8	8.3	10.9	3.5	2.8	2.4	4.5	4.3	6.4	7.8	6.7	6.2
Pseudomonas	4.0	3.7	4.3	4.4	19.9[a]	12.4[b]	5.3[b]	4.0[b]	8.5[a]	6.1[a]	0.9[b]	0.4[b]	1.3	1.1	1.1	0.9
Acinetobacter	5.8	3.9	6.3	6.7	2.6	1.4	0.9	1.2	6.0	4.7	6.0	1.0	4.9	3.1	2.2	3.0
Cellulomonas	0.1	0.1	0.3	0.2	0.1	0.1	0.0	0.1	2.8	6.1	8.0	5.5	5.9	5.0	3.8	7.5
Bacteroides	3.1	3.5	2.6	5.6	1.0	1.4	1.2	1.1	2.3	2.5	2.2	1.5	3.7	3.9	4.8	3.9
Escherichia	4.7	4.0	3.9	3.3	0.5	0.9	0.3	0.5	4.5[a]	3.3[a]	0.2[b]	0.1[b]	0.3	0.3	0.3	0.1
Corynebacterium	2.4	2.4	3.1	1.9	0.4	0.3	0.2	0.4	2.9	2.1	1.0	0.4	2.6	1.9	1.9	2.2
Fusobacterium	1.2	1.3	0.8	1.5	0.6	0.7	0.8	0.5	0.9	2.4	2.8	0.9	1.6	1.7	1.1	0.5
Pasteurella	1.5	0.2	0.3	0.2	3.3[a]	1.2[b]	2.6[ab]	0.7[ab]	1.4	0.4	0.7	0.4	0.4	0.9	0.4	0.2
Pedobacter	1.1	1.0	1.2	1.7	0.2	0.2	0.1	0.1	2.6	2.8	1.2	1.6	1.1	1.1	1.6	1.3
Streptococcus	1.9	1.9	1.6	1.5	1.4	1.1	2.2	2.1	0.5	0.9	0.7	0.4	0.5	0.5	0.5	0.6
Serratia	3.0	2.5	2.4	2.2	0.4	0.5	0.2	0.3	2.2	1.7	0.1	0.1	0.3	0.2	0.2	0.1
Prevotella	0.9	1.1	0.7	1.6	0.4	0.5	0.2	0.6	1.1	0.7	0.5	0.2	1.8	1.4	2.6	1.7
Staphylococcus	1.3	0.9	1.2	1.1	0.2	0.2	0.1	0.1	2.3	2.1	0.7	0.2	0.7	0.7	0.8	0.6
Ruminococcus	1.2	1.4	1.4	1.8	0.3	0.4	0.1	0.3	0.8	0.6	0.7	0.4	0.9	1.1	1.7	1.6
Candidatus Blochmannia	0.9	0.9	0.9	1.0	0.2	0.3	0.2	0.4	1.5	1.1	0.7	0.6	1.1	1.0	1.3	1.1
Porphyromonas	1.2	1.6	1.1	1.8	0.1	0.3	0.1	0.3	0.5	0.4	0.3	0.2	1.2	1.5	1.5	1.3

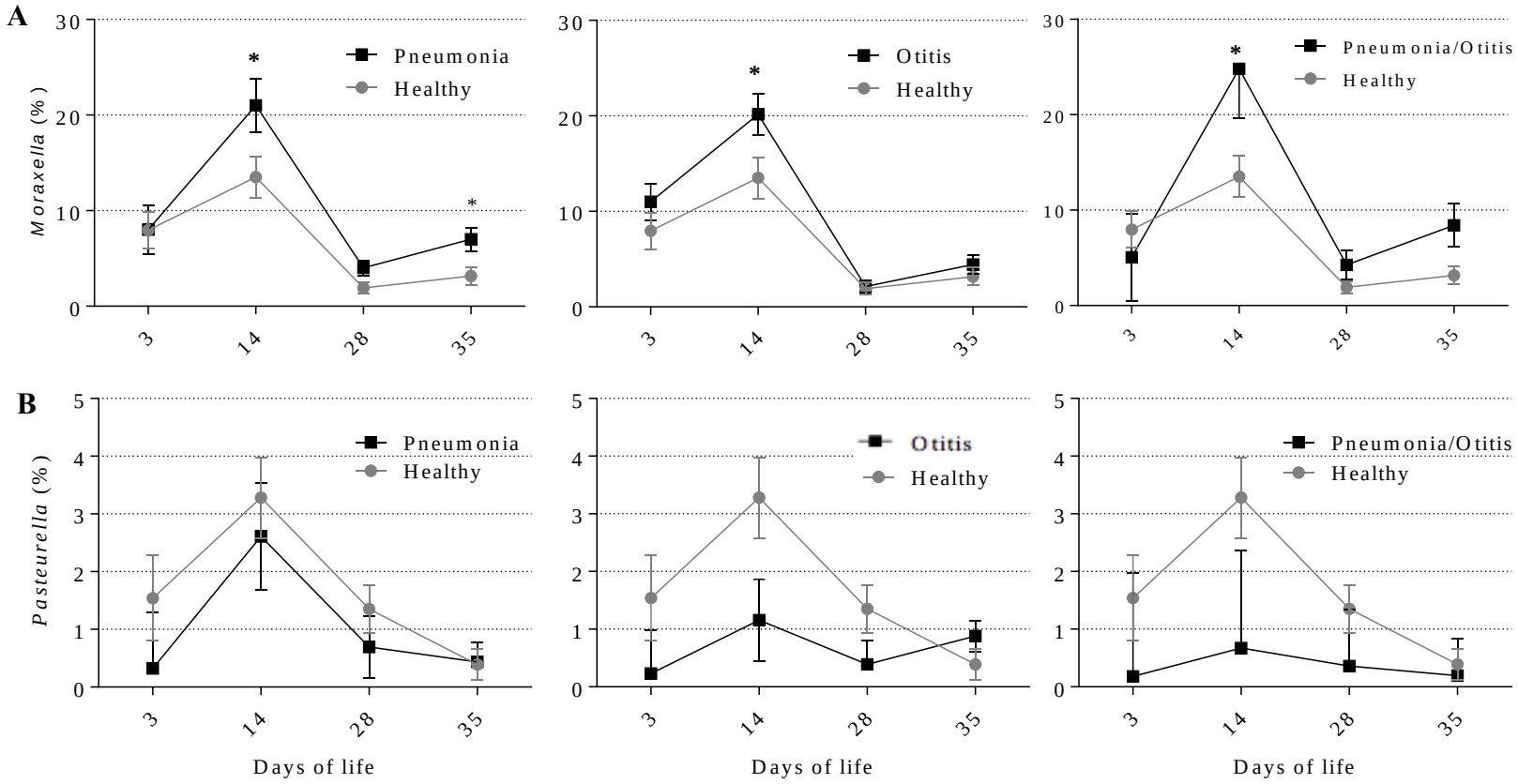
[1]H = Healthy; [2]O = Otitis; [3]P = Pneumonia; [4]P&O = Pneumonia and otitis combined

Figure 4.4: Mean relative abundance of the genus *Mannheimia* (A) and *Mycoplasma* (B) according to postnatal age at sample collection (3, 14, 28, 35 days) and health status (healthy, otitis, pneumonia, and pneumonia-otitis combined). Error bars are positioned around the means and represent the standard error of the mean. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the status "healthy" within each sample collection time point. Asterisks on a series of data points indicate a significant difference ($P < 0.05$) between the respective health status categories within postnatal age.



The relative abundance of each of *Mannheimia* and *Moraxella* at day 14 in calves diagnosed with pneumonia was significantly higher when compared to healthy calves (Figure 4.4A and 4.5A). Similar results were observed for *Mannheimia* and *Mycoplasma* at day 28 (Figure 4.4A and 4.4B). The relative abundance of *Pasteurella* at day 14 was lower in calves diagnosed with otitis when compared to healthy calves (Figure 4.5B). The genera *Pseudomonas*, *Escherichia*, and *Corynebacterium* were also subjected to more detailed analyses and these are presented in supplemental material 2, 3, and 4.

Figure 4.5: Mean relative abundance of the genus *Moraxella* (A) and *Pasteurella* (B) according to postnatal age at sample collection (3, 14, 28, 35 days) and health status (healthy, otitis, pneumonia, and pneumonia-otitis combined). Error bars are positioned around the means and represent the standard error of the mean. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the status "healthy" within each sample collection time point. Asterisks on a series of data points indicate a significant difference ($P < 0.05$) between the respective health status categories within postnatal age.



DISCUSSION

To our knowledge, this is the first report to have longitudinally evaluated the URT microbiota of Holstein dairy calves from birth (day 3 of life) until 35 days of life. The URT is a critical point of entry for pathogens, and thus a potential route for infection of the lower respiratory tract (Laurenzi et al., 1961) and the middle ear (Murphy et al., 2009). Therefore, characterization of the URT is a crucial step in unraveling the pattern of development of both pneumonia and infection of the middle ear in calves. Here, we reported that the genera *Mannheimia, Moraxella*, and *Mycoplasma* were found in significantly higher abundances in dairy calves that developed pneumonia, otitis, or pneumonia-otitis combined during the pre-weaned period. We also observed that calves affected with pneumonia, otitis and pneumonia-otitis combined at day 3 of life presented the same bacterial community structure when compared to healthy animals, however animals diagnosed with pneumonia only at this same age had a significantly higher bacterial load, as defined by the log of the copy numbers of the 16S rRNA gene quantified by quantitative PCR technique, soon after they were born, than healthy calves. Together, these results suggest that the microbial composition of the URT of newborn calves (3 days of life) was not predictive of pneumonia, but the total URT bacterial load was. Therefore, pneumonia and otitis pathogens are already present in the URT of newborn calves, but those that will eventually develop respiratory disease simply have a higher load of total bacteria at 3 days of life.

The incidence of pneumonia and otitis reported in the present study was higher than the rates reported by the NAHMS 2002 (USDA, 2002) and 2007 (USDA, 2007); the reported incidence of BRD was 12.4% versus the 30% incidence herein reported. Our study was conducted on a commercial dairy farm that used a group house system (20 calves per group) and

calves were fed ad-libitum acidified non-saleable milk. These types of systems have been reported to be associated with higher incidence of respiratory diseases such as pneumonia (Anderson et al., 2008). Svensson et al. (2003) evaluated the two types of pre-weaning housing systems, in which calves were raised in individual pens and milk was fed manually, or calves raised in grouping pens with automatic milk-feeding system (Svensson et al., 2003). Still in Svensson study, a higher odds ratios for respiratory disease and also increased respiratory sounds in calves housed in group pens with an automatic milk-feeding system (OR: 2.2, 2.8) than calves housed in individually pens was observed. Additionally, the increased chance of transmission of pathogenic agents between calves in housing group systems is also observed, since calves in this type of system tend to be more densely housed, resulting in closer animal-to-animal contact, and consequently propagation of the infections (Steenkamer et al., 1982, Maatje et al., 1993, Hepola, 2003). Therefore, this justifies the higher incidence rates reported in this present study.

In this present study, all calves that were diagnosed with pneumonia received antibiotic therapy. Systemic antibiotic therapy in calves has been reported to alter the fecal microbiome (Oikonomou et al., 2013) and most likely should also impact the microbiome of the URT. Therefore, the URT microbiome of post disease diagnostic here described has most likely been altered by the use of systemic antibiotics. Our study was conducted on a commercial dairy farm and farm protocols could not be modified solely for the purpose of the study. Additionally, pneumonia and otitis are diseases commonly caused by bacterial infection and neglecting to treat sick calves with the proper antibiotic therapy could be considered inhumane and may not have been approved by the Cornell Institutional Animal Care and Use Committee. Nevertheless, this is a limitation of the current study and future work should explore the URT microbiome for calves affected or not by respiratory diseases and treated or not by systemic antibiotic therapy.

An important result of the present study was the association between the genus *Moraxella* and, disease statuses. Calves that were diagnosed with pneumonia, otitis, or pneumonia-otitis combined had a significantly higher abundance of *Moraxella* at day 14 when compared to healthy calves. Although members of the genus *Moraxella* are often isolated from cases of keratoconjunctivitis, an important ocular illness in bovines (Angelos et al., 2007, Peek, 2008, Galvão and Angelos, 2010), there are few reports in the scientific literature describing an association of the genus *Moraxella* with pneumonia and/or otitis media. Catry et al. (2007) isolated *Moraxella ovis* from the upper and lower respiratory tract of calves affected with acute and chronic respiratory disease (Catry, 2007), and Corbeil et al. (1985) reported that *Moraxella* spp. enhance the growth of RT bacterial pathogens such as *Mannheimia haemolytica*, *Pasteurella multocida*, and *Haemophilus somnus* (Corbeil et al., 1985). The association between the genus *Moraxella* and the incidence of pneumonia and otitis in pre-weaned dairy cattle described in the present study is novel and of potential significance.

The abundance of *Mannheimia* and *Mycoplasma* increased substantially over time. At days 14, *Mannheimia* were significantly more abundant in disease statuses, and the same trend was observed for *Mannheimia* and *Mycoplasma* at day 28 of life. It is known that *Mannheimia haemolytica* is a commensal organism that inhabits the nasopharynx and can lead to disease when calves are exposed to stress factors such as weaning, comingling, and coinfection with others microorganisms (Rice et al., 2007), such as *Mycoplasma bovis* (Maunsell and Donovan, 2009, Griffin et al., 2010), which is suggested by ours results. Furthermore, *Pasteurella*, another important bacterium for the BRD complex in calves (Thomson et al., 1975, Allen, 1991, Fulton et al., 2009), was not associated with disease nor was the genus highly abundant in any health status or age group in our analyses. Our data are supported by the results of Klima et al., who

detected a lower occurrence of *Pasteurella multocida* in a BRD outbreak in North American feedlots (Klima, 2014).

In our study the most abundant genera detected in the URT of dairy calves were *Mannheimia, Mycoplasma, Moraxella, Psychrobacter,* and *Pseudomonas*. Recently, Holman et al. (2015) investigated the nasopharyngeal bacterial community of feedlot cattle at the day of feedlot entry and also at 60 days after entrance (Holman et al., 2015). The Holman experiment focused on differences in the microbial community over time, and comparisons between healthy and diseased animals were not evaluated. Still, in the Holman study, *Staphylococcus*, *Mycoplasma, Mannheimia,* and *Moraxella* were the dominant genera identified in the URT samples collected at 60 days after feedlot arrival, comparable to what we observed in young calves with bacterial infections. The Holman results at 60 days may in part be a consequence of the high levels of stress (caused by shipping, change of environment, and stocking density) (Snowder et al., 2006) that beef calves experience following their arrival on the feedlot.

The genera *Mycoplasma, Mannheimia, Pasteurella, Staphylococcus,* and *Streptococcus* were detected in all animals at four time points examined, regardless of health conditions. These pathogens have been repeatedly described as the primary cause of otitis media (Walz, 1997, Maunsell, 2012) and/or pneumonia (Angen, 2009, Pardon, 2011, Maunsell, 2012). A potential explanation for this finding could be that disease develops when host and/or pathogen factors result in bacterial proliferation and dissemination to other body sites (Maunsell and Donovan, 2009), and/or as a result of a detrimental host inflammatory response (Faden et al., 1995, García-Rodríguez and Fresnadillo Martínez, 2002). Further studies are required to examine both the microbial composition of the RT and the relative contribution of the immune system to the calf's RT health.

The microbial diversity is understood as being a function of the number of different categories (richness) and the relative distribution of individual elements among these categories (evenness) (Washington, 1984). Shannon diversity and Chao 1 richness indexes were not significantly different between healthy calves and calves diagnosed with BRD. The high microbial diversity detected in our experiment was expected since the URT is the first compartment of the respiratory system that is in close contact with the environment. In agreement with our findings, Charlson et al. (2011) performed an intensive sampling of multiple sites along the respiratory tract of healthy human individuals and observed low levels of bacterial diversity in the lower respiratory tract, but high levels of bacterial diversity in the URT (Charlson et al., 2011). Additionally, Huang et al. (2010) detected high diversity levels in the airway bacterial community of patients with chronic obstructive pulmonary disease (Huang et al., 2010). Interesting a negative correlation was observed between the Shannon diversity index and bacterial load, indicating that when microbial colonization increases there is a corresponding reduction in microbial diversity. This finding is supported by Boutin et al. (2014), who reported a negative correlation between bacterial load and alpha-diversity index in the nasal cavity of human (Boutin, 2015)."

CONCLUSIONS

In summary, this study demonstrated that the URT of Holstein dairy calves from 3 days to 35 days of life encompasses a highly rich and diverse bacterial community. Thirty genera were shared between all ages and health statuses and *Mannheimia, Mycoplasma, Moraxella* were the three most common bacterial genera detected in the calves URT. Our data supports the concept that *Mannheimia* and *Mycoplasma* are two dominant bacteria associated with pneumonia and

otitis. Additionally, the genus *Moraxella* could play an important role in the pathogenesis of pneumonia and otitis, and the high neonatal bacterial load is a significant contributor to the development of pneumonia. These results provide an unprecedented understanding of the evolution of the bovine URT microbiome in pre-weaning calves and its association with RT diseases.

REFERENCES

Allen, J. W., Viel, L., Bateman, K. G., Rosendal, S., Shewen, P. E., Physick-Sheard, P. 1991. The microbial flora of the respiratory tract in feedlot calves: associations between nasopharyngeal and bronchoalveolar lavage cultures. Can J Vet Res 55(4):341-346.

Amann, R. I., W. Ludwig, and K. H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev 59(1):143-169.

Anderson, N. G. 2008. Experiences with free-access acidified-milk feeding in Ontario. Pages 12-24. *Stillwater (OK): Frontier Printers Inc, Proceedings of the 41st Annual Meeting of the American Association of Bovine Practitioners, Charlotte (NC)*.

Angelos, J. A., P. Q. Spinks, L. M. Ball, and L. W. George. 2007. *Moraxella bovoculi* sp. nov., isolated from calves with infectious bovine keratoconjunctivitis. Int J Syst Evol Microbiol 57(Pt 4):789-795.

Angen, O., Thomsen, J., Larsen, L. E., Larsen, J., Kokotovic, B., Heegaard, P. M., Enemark, J. M. 2009. Respiratory disease in calves: microbiological investigations on trans-tracheally aspirated bronchoalveolar fluid and acute phase protein response. Vet Microbiol 137(1-2):165-171.

Arcangioli, M. A., H. Aslan, F. Tardy, F. Poumarat, and D. Le Grand. 2012. The use of pulsed-field gel electrophoresis to investigate the epidemiology of *Mycoplasma bovis* in French calf feedlots. Vet J 192(1):96-100.

Baba, A. I., O. Rotaru, and G. Rapuntean. 1988. Middle ear infection in suckling and weaned calves. Morphol Embryol (Bucur) 34(4):271-275.

Babiuk, L. A., M. J. Lawman, and H. B. Ohmann. 1988. Viral-bacterial synergistic interaction in respiratory disease. Adv Virus Res 35:219-249.

Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., Mills, D. A., Caporaso, J. G. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Methods 10(1):57-59.

Bosch, A. A., G. Biesbroek, K. Trzcinski, E. A. Sanders, and D. Bogaert. 2013. Viral and bacterial interactions in the upper respiratory tract. PLoS Pathog 9(1):e1003057.

Boutin, S., Graeber, S. Y., Weitnauer, M., Panitz, J., Stahl, M., Clausznitzer, D., Kaderali, L., Einarsson, G., Tunney, M. M., Elborn, J. S., Mall, M. A., Dalpke, A. H. 2015. Comparison of microbiomes from different niches of upper and lower airways in children and adolescents with cystic fibrosis. PLoS One 10(1):e0116029.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., Knight, R. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6(8):1621-1624.

Catry, B., Boyen, F., Baele, M., Dewulf, J., de Kruif, A., Vaneechoutte, M., Haesebrouck, F., Decostere, A. 2007. Recovery of Moraxella ovis from the bovine respiratory tract and differentiation of *Moraxella* species by tDNA-intergenic spacer PCR. Vet Microbiol 120(3-4):375-380.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. Pages 265-270. Scandinavian Journal of Statistics, Theory and Applications.

Charlson, E. S., K. Bittinger, A. R. Haas, A. S. Fitzgerald, I. Frank, A. Yadav, F. D. Bushman, and R. G. Collman. 2011. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med 184(8):957-963.

Chase, C. C., Hurley, D. J., Reber, A. J. 2008. Neonatal immune development in the calf and its impact on vaccine response. Vet Clin North Am Food Anim Pract 24(1):87-104.

Confer, A. W. 2009. Update on bacterial pathogenesis in BRD. Anim Health Res Rev 10(2):145-148.

Corbeil, L. B., W. Woodward, A. C. Ward, W. D. Mickelsen, and L. Paisley. 1985. Bacterial interactions in bovine respiratory and reproductive infections. J Clin Microbiol 21(5):803-807.

Czuprynski, C. J., F. Leite, M. Sylte, C. Kuckleburg, R. Schultz, T. Inzana, E. Behling-Kelly, and L. Corbeil. 2004. Complexities of the pathogenesis of *Mannheimia haemolytica* and *Haemophilus somnus* infections: challenges and potential opportunities for prevention? Anim Health Res Rev 5(2):277-282.

D'Argenio, V. and F. Salvatore. 2015. The role of the gut microbiome in the healthy adult status. Clin Chim Acta 451(Pt A):97-102.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460-2461.

Faden, H., L. Duffy, A. Williams, D. A. Krystofik, and J. Wolf. 1995. Epidemiology of nasopharyngeal colonization with nontypeable Haemophilus influenzae in the first 2 years of life. J Infect Dis 172(1):132-135.

Francoz, D., G. Fecteau, A. Desrochers, and M. Fortin. 2004. Otitis media in dairy calves: a retrospective study of 15 cases (1987 to 2002). Can Vet J 45(8):661-666.

Fulton, R. W., Blood, K. S., R. J. Panciera, M. E. Payton, J. F. Ridpath, A. W. Confer, J. T. Saliki, L. T. Burge, R. D. Welsh, B. J. Johnson, and A. Reck. 2009. Lung pathology and

infectious agents in fatal feedlot pneumonias and relationship with mortality, disease onset, and treatments. J Vet Diagn Invest 21(4):464-477.

Fulton, R. W., Cook, B. J., Step, D. L., Confer, A. W., Saliki, J. T., Payton, M. E., Burge, L. J., Welsh, R. D., Blood, K. S. 2002. Evaluation of health status of calves and the impact on feedlot performance: assessment of a retained ownership program for postweaning calves. Can J Vet Res 66(3):173-180.

Galvão, K. N. and J. A. Angelos. 2010. Ulcerative blepharitis and conjunctivitis in adult dairy cows and association with *Moraxella bovoculi*. Can Vet J 51(4):400-402.

García-Rodríguez, J. A. and M. J. Fresnadillo Martínez. 2002. Dynamics of nasopharyngeal colonization by potential respiratory pathogens. J Antimicrob Chemother 50 Suppl S2:59-73.

Gilbert, J. A., Meyer, F., Antonopoulos, D., Balaji, P., Brown, C. T., Desai, N., Eisen, J. A., Evers, D., Field, D., W. Feng, Huson, D., Jansson, J., Knight, R., Knight, J., Kolker, E., Konstantindis, K., Kostka, J., Kyrpides, N., and R. Mackelprang, McHardy, A., Quince, C., Raes, J., Sczyrba, A., Shade, A., Stevens, R. 2010. Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. Stand Genomic Sci 3(3):243-248.

Gorden, P. J. and P. Plummer. 2010. Control, management, and prevention of bovine respiratory disease in dairy calves and cows. Vet Clin North Am Food Anim Pract 26(2):243-259.

Griffin, D., M. M. Chengappa, J. Kuszak, and D. S. McVey. 2010. Bacterial pathogens of the bovine respiratory disease complex. Vet Clin North Am Food Anim Pract 26(2):381-394.

Haines, D. M., K. M. Martin, E. G. Clark, G. K. Jim, and E. D. Janzen. 2001. The immunohistochemical detection of *Mycoplasma bovis* and bovine viral diarrhea virus in

tissues of feedlot cattle with chronic, unresponsive respiratory disease and/or arthritis. Can Vet J 42(11):857-860.

Hepola, H. 2003. Milk feeding systems for dairy calves in groups: Efects on feed intakes, growth and health. Pages 233-243. Applied Animal Behavior Science.

Holman, D. B., T. A. McAllister, E. Topp, A. D. Wright, and T. W. Alexander. 2015. The nasopharyngeal microbiota of feedlot cattle that develop bovine respiratory disease. Vet Microbiol 180(1-2):90-95.

Huang, Y. J., E. Kim, M. J. Cox, E. L. Brodie, R. Brown, J. P. Wiener-Kronish, and S. V. Lynch. 2010. A persistent and diverse airway microbiota present during chronic obstructive pulmonary disease exacerbations. OMICS 14(1):9-59.

Jensen, R., Maki, L. R., Lauerman, L. H., Raths, W. R., Swift, B. L., Flack, D. E., Hoff, R. L., Hancock, H. A.,, Tucker, J. O., Horton, D. P., Weibel, J. L. 1983. Cause and pathogenesis of middle ear infection in young feedlot cattle. J Am Vet Med Assoc 182(9):967-972.

Khafipour, E., S. Li, J. C. Plaizier, and D. O. Krause. 2009. Rumen microbiome composition determined using two nutritional models of subacute ruminal acidosis. Appl Environ Microbiol 75(22):7115-7124.

Klima, C. L., Zaheer, R., Cook, S. R., Booker, C. W., Hendrick, S., Alexander, T. W., McAllister, T. A. 2014. Pathogens of bovine respiratory disease in North American feedlots conferring multidrug resistance via integrative conjugative elements. J Clin Microbiol 52(2):438-448.

Kolbert, C. P. and D. H. Persing. 1999. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. Curr Opin Microbiol 2(3):299-305.

Lago, A., McGuirk, S. M., Bennett, T. B., Cook, N. B., Nordlund, K. V. 2006. Calf respiratory disease and pen microenvironments in naturally ventilated calf barns in winter. J Dairy Sci 89(10):4014-4025.

Laurenzi, G. A., R. T. Potter, and E. H. Kass. 1961. Bacteriologic flora of the lower respiratory tract. N Engl J Med 265:1273-1278.

Maatje, K., J. Verhoeff, W. D. Kremer, A. L. Cruijsen, and T. S. van den Ingh. 1993. Automated feeding of milk replacer and health control of group-housed veal calves. Vet Rec 133(11):266-270.

Maeda, T., Shibahara, T., Kimura, K., Wada, Y., Sato, K., Imada, Y., Ishikawa, Y., Kadota, K. 2003. *Mycoplasma bovis*-associated suppurative otitis media and pneumonia in bull calves. J Comp Pathol 129(2-3):100-110.

Maunsell, F., Brown, M. B., Powe, J., Ivey, J., Woolard, M., Love, W., Simecka, J. W. 2012. Oral inoculation of young dairy calves with Mycoplasma bovis results in colonization of tonsils, development of otitis media and local immunity. PLoS One 7(9):e44523.

Maunsell, F. P. and G. A. Donovan. 2009. Mycoplasma bovis Infections in young calves. Vet Clin North Am Food Anim Pract 25(1):139-177, vii.

McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., Andersen, G. L., Knight, R., Hugenholtz, P. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 6(3):610-618.

McEwen, S. A. and T. J. Hulland. 1985. *Haemophilus somnus*-induced Otitis and Meningitis in a Heifer. Can Vet J 26(1):7-8.

Murphy, T. F., L. O. Bakaletz, and P. R. Smeesters. 2009. Microbial interactions in the respiratory tract. Pediatr Infect Dis J 28(10 Suppl):S121-126.

Nonnenmacher, C., A. Dalpke, R. Mutters, and K. Heeg. 2004. Quantitative detection of periodontopathogens by real-time PCR. J Microbiol Methods 59(1):117-125.

Oikonomou, G., V. S. Machado, C. Santisteban, Y. H. Schukken, and R. C. Bicalho. 2012. Microbial diversity of bovine mastitic milk as described by pyrosequencing of metagenomic 16s rDNA. PLoS One 7(10):e47671.

Oikonomou, G., A. G. Teixeira, C. Foditsch, M. L. Bicalho, V. S. Machado, and R. C. Bicalho. 2013. Fecal microbial diversity in pre-weaned dairy calves as described by pyrosequencing of metagenomic 16S rDNA. Associations of Faecalibacterium species with health and growth. PLoS One 8(4):e63157.

Pardon, B., De Bleecker, K., Dewulf, J., Callens, J., Boyen, F., Catry, B., Deprez, P. 2011. Prevalence of respiratory pathogens in diseased, non-vaccinated, routinely medicated veal calves. Vet Rec 169(11):278.

Pardon, B., Hostens, M., Duchateau, L., Dewulf, J., De Bleecker, K., Deprez, P. 2013. Impact of respiratory disease, diarrhea, otitis and arthritis on mortality and carcass traits in white veal calves. BMC Vet Res 9:79.

Peek, T. J. D. a. S. F. 2008. Diseases of dairy cattle. Vol. Second edition. Saunders Elsevier.

Rice, J. A., L. Carrasco-Medina, D. C. Hodgins, and P. E. Shewen. 2007. *Mannheimia haemolytica* and bovine respiratory disease. Anim Health Res Rev 8(2):117-128.

Schloss, P. D. and J. Handelsman. 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl Environ Microbiol 71(3):1501-1506.

Shahriar, F. M., E. G. Clark, E. Janzen, K. West, and G. Wobeser. 2002. Coinfection with bovine viral diarrhea virus and *Mycoplasma bovis* in feedlot cattle with chronic pneumonia. Can Vet J 43(11):863-868.

Snowder, G. D., L. D. Van Vleck, L. V. Cundiff, and G. L. Bennett. 2006. Bovine respiratory disease in feedlot cattle: environmental, genetic, and economic factors. J Anim Sci 84(8):1999-2008.

Steenkamer, N. 1982. Alternative housing systems for veal calves, their effect on welfare and performance and their economic feasibility. . Pages 226-234. Vol. 19. Signoret, J.P., Welfare and Husbandry of Calves: Current Topics in Veterinary Medicine and Animal Science.

Svensson, C., K. Lundborg, U. Emanuelson, and S. O. Olsson. 2003. Morbidity in Swedish dairy calves from birth to 90 days of age and individual calf-level risk factors for infectious diseases. Prev Vet Med 58(3-4):179-197.

Thomson, R. G., S. Chander, M. Savan, and M. L. Fox. 1975. Investigation of factors of probable significance in the pathogenesis of pneumonic pasteurellosis in cattle. Can J Comp Med 39(2):194-207.

USDA. 2002. Part I: Reference of dairy health and management practices in the United States. in FortCollins (CO): USDA: APHIS: VS, CEAH, National Animal Health Monitoring System. #N377.1202 ed.

USDA. 2007. Heifer Calf Health and Management Practices on U.S. Agriculture, ed. USDA:APHIS:VS:CEAH NRRC Building B, M.S. 2E7 2150 Centre Avenue Fort Collins, CO 80526-811.

Van Biervliet, J., Perkins, G. A., Woodie, B., Pelligrini-Massini, A., Divers, T. J., de Lahunta, A. 2004. Clinical signs, computed tomographic imaging, and management of chronic otitis media/interna in dairy calves. J Vet Intern Med 18(6):907-910.

Virtala, A. M., Mechor, G. D., Gröhn, Y. T., Erb, H. N. 1996. The effect of calfhood diseases on growth of female dairy calves during the first 3 months of life in New York State. J Dairy Sci 79(6):1040-1049.

Waltner-Toews, D., Martin, S. W., Meek, A. H. 1986. The effect of early calfhood health status on survivorship and age at first calving. Can J Vet Res 50(3):314-317.

Walz, P. H., Mullaney, T. P., Render, J. A., Walker, R. D., Mosser, T., Baker, J. C. 1997. Otitis media in preweaned Holstein dairy calves in Michigan due to *Mycoplasma bovis*. J Vet Diagn Invest 9(3):250-254.

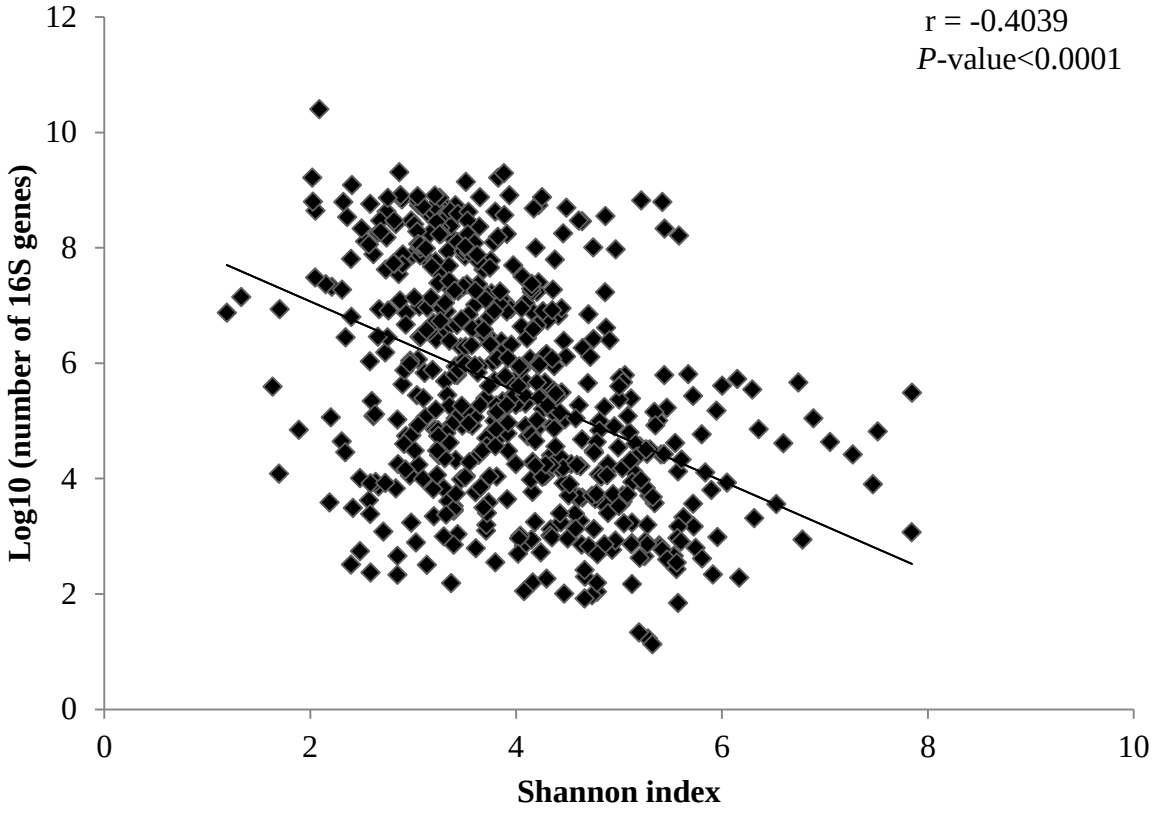
Warnick, L. D., Erb, H. N., White, M. E. 1995. Lack of association between calf morbidity and subsequent first lactation milk production in 25 New York Holstein herds. J Dairy Sci 78(12):2819-2830.

Washington, H. G. 1984. Diversity, biotic and similarity indices. Pages 653-694. Water Res.

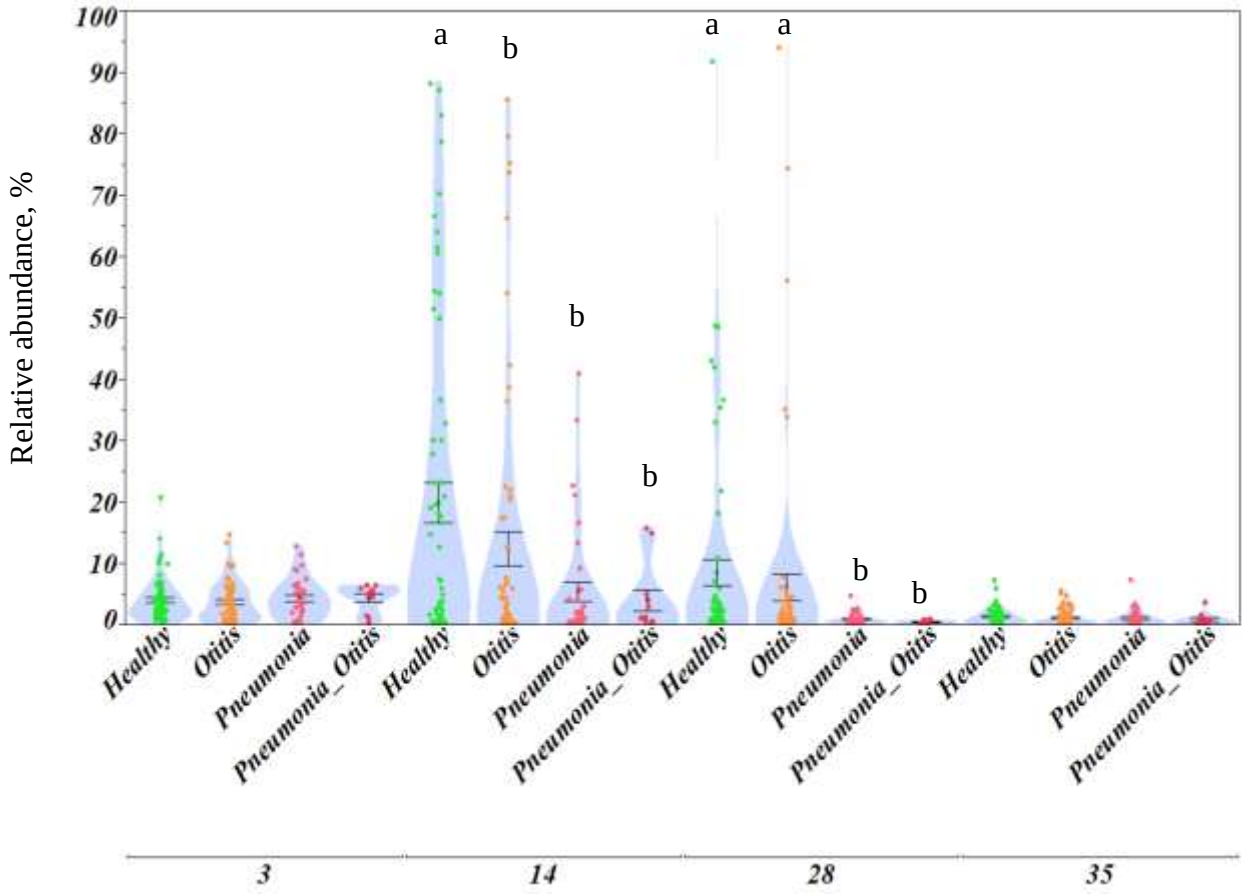
Webster, A. J. 1983. Environmental stress and the physiology, performance and health of ruminants. J Anim Sci 57(6):1584-1593.

Yeruham, I., D. Elad, and M. Liberboim. 1999. Clinical and microbiological study of an otitis media outbreak in calves in a dairy herd. Zentralbl Veterinarmed B 46(3):145-150.

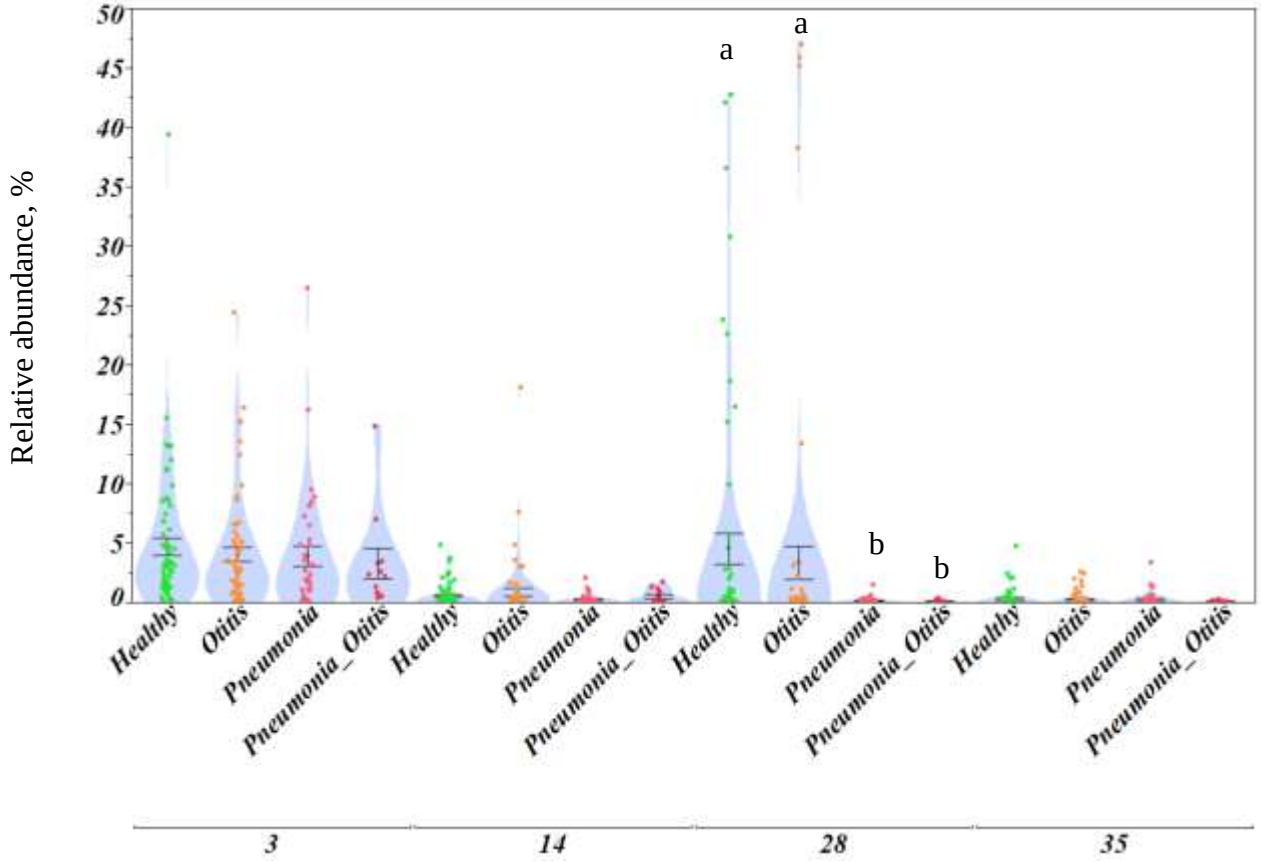
Supplemental 4.1: Correlation between alpha-diversity and bacterial load of the neonatal calf upper respiratory tract. The alpha-diversity is represented by the Shannon evenness index. The bacterial load was measured via proxy of the number of 16S rRNA genes. Samples from all sampling time points are represented.



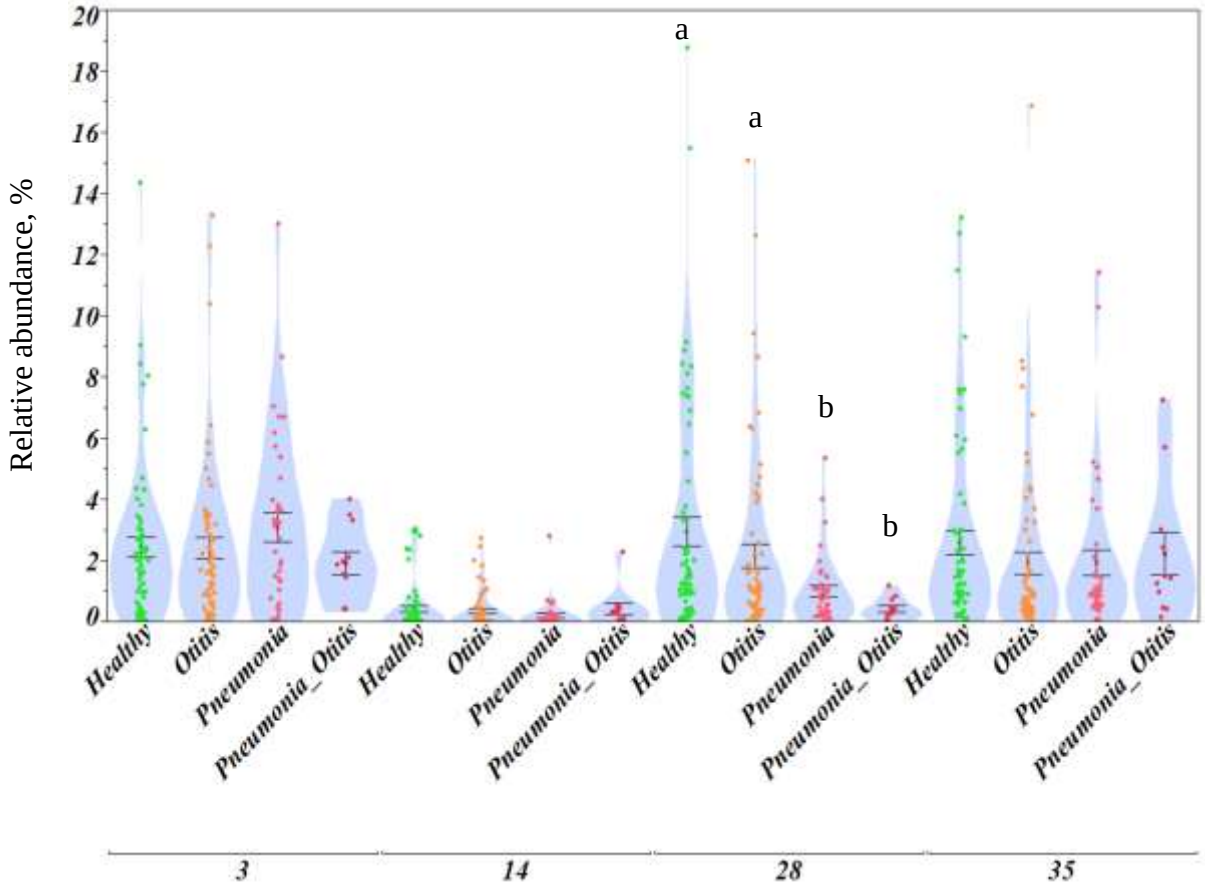
Supplemental 4.2: Mean relative abundance of the genus *Pseudomonas* according to postnatal age at sample collection (3, 14, 28, 35 days) and health status (healthy = green, otitis = orange, pneumonia = pink, and pneumonia-otitis combined = red). Error bars are positioned around the means and represent the standard error of the mean, and contours show regions of data density. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia, and pneumonia-otitis combined) against the status "healthy" within each sample collection age. [a,b] different superscripts on a series of data points indicate a significant difference ($P < 0.05$) between the respective health status categories within each age group sampled.



Supplemental 4.3: Mean relative abundance of the genus *Escherichia* according to postnatal age at sample collection (3, 14, 28, 35 days) and health status (healthy = green, otitis = orange, pneumonia = pink, and pneumonia-otitis combined = red). Error bars are positioned around the means and represent the standard error of the mean, and contours show regions of data density. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia and pneumonia-otitis combined) against the status "healthy" within each sample collection age. [a,b] different superscripts on a series of data points indicate a significant difference ($P < 0.05$) between the respective health status categories within each age group sampled.



Supplemental 4.4: Relative abundance of the genus *Corynebacterium* by health status (healthy = green, otitis = orange, pneumonia = light red, and pneumonia-otitis = dark red) and postnatal age at sample collection (3, 14, 28, and 35 days). Error bars are positioned around the means and represent the standard error of the mean, and contours show regions of data density. Dunnett's multiple comparison procedure was used to compare each disease status (otitis, pneumonia and pneumonia-otitis combined) against the status "healthy" within each sample collection age. [a,b] different superscripts on a series of data points indicate a significant difference ($P < 0.05$) between the respective health status categories within each age group sampled.



CHAPTER FIVE

THE MATERNAL MICROBIOME PLAYS A KEY ROLE IN DETERMINING THE OFFSPRING'S EARLY-LIFE MICROBIAL COMMUNITY OF *BOS TAURUS*

Svetlana Ferreira Lima[*], Marcela Lucas de Souza Bicalho[*], Rodrigo Carvalho Bicalho[*1]

[*]Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

ABSTRACT

Natural transference of maternal microbes to the neonate, especially at birth via the vaginal canal, has recently been recognized in humans; however, the same process has not yet been documented in bovines. Therefore, our study was designed to elucidate the potential vertical transfer of bacterial organisms from cows to their neonatal calves and investigate its microbial influence on calf health during the pre-weaning period. Here, by applying deep sequencing, we compared the bacterial communities in vaginal and fecal samples from 81 pregnant Holstein dairy cows (≥ 273 days carried calf) from a single farm versus those in nasopharyngeal and fecal samples collected at 3, 14 and 35 days of life from their respective progeny. Multivariate analysis with Unifrac revealed unique profiles of bacterial communities by sample type, wherein the dissimilarities between sample types were mainly accounted for by the presence or absence of operational taxonomic units (OTUs) rather than by their relative abundances. However, the microbiota of the calf upper respiratory tract (URT), regardless of calf age, was found to be highly similar to the maternal vaginal microbiota. Moreover, the calf fecal microbiota clustered closely to the maternal fecal microbiota, progressing toward an adult-like state over the first 35 days of life when only relative abundances of taxa were considered. Sixty four, 65 and 87% of the detected OTUs were shared between cow and calf fecal microbiota at days 3, 14 and 35 of life respectively, whereas 73, 76 and 87% of detected OTUs were shared between the maternal vaginal microbiome and the calf URT microbiota at days 3, 14 and 35 of life, respectively. *Bacteroidetes*, *Ruminococcus*, *Clostridium*, and *Blautia* were the top four genera identified in dam and calf fecal samples, regardless of calf age. *Mannheimia, Moraxella*, *Bacteroides, Streptococcus* and *Pseudomonas* were the top five genera identified within the most abundant bacterial genera in dam vaginal and calf URT samples across all days of calf life

examined. Furthermore, the genus *Mannheimia* was relatively more abundant in the vaginal microbiota of dams whose progeny were diagnosed with respiratory and middle ear disease. Our results indicate that the dam vaginal microbiota potentially influences the initial bacterial colonization of the calf URT, and that it may have a beneficial effect on the health of the calf respiratory tract and middle ear.

INTRODUCTION

Our understanding of the complex ruminant microbiome and recognition of the importance of the host-microbiome interaction have significantly expanded in the last few years. Bovine microbial communities have been characterized in depth across the gastrointestinal tract (Dill-McFarland et al, 2017; Jami et al, 2013; Meale et al, 2016) and at other anatomical sites, including the mammary gland (Addis et al, 2016), uterus (Machado et al, 2012; Santos and Bicalho 2012), and airways (Holman et al, 2015). It has been shown that the composition of the bovine microbiome can affect host health status (Lima et al, 2016; Oikonomou et al, 2014; Zinicola et al, 2015) and animal performance (Lima et al, 2015). However, despite these impressive advances, there is still limited information on how and when the microbiome begins to be established in the bovine host.

Studies in humans unveiled the potential for transference of bacterial communities from the mother to her progeny during pregnancy (Aagaard et al, 2014; Gosalbes et al 2013) and the vertical transference via the birth canal during delivery (Dominguez-Bello et al, 2010; Lif Holgerson et al, 2011). Numerous immunological and hormonal changes during pregnancy affect maternal immune regulation, leading to an anti-inflammatory biased response. This phenomenon also affects the developing immune system of the fetus (Chase et al, 2008; Morein et al, 2002;

Ridge et al, 1996), which might facilitate perinatal microbial colonization, especially at birth, by the mother's microbiome. Exposure to microbes early in life has been shown to influence maturation of the neonatal immune system (Chung et al, 2012; Kelly et al, 2007) and potentially its metabolic development (Cho et al, 2012; Cox et al, 2014). Furthermore, its disruption, by interrupting the vertical transference of bacterial organisms from the mother to the newborn (e.g., as occurs with the C-section procedure), has been shown to be associated with negative downstream consequences such as asthma and allergic disorders (Renz-Polster et al, 2005; Salam et al, 2006). In veterinary medicine, investigation of the bovine female microbiome as a factor that influences microbial colonization early in the life of dairy calves, and consequently as a contributor to neonatal calf health and/or disease predisposition, is still underexplored.

The major cause of pre-weaning and weaning calf mortality in the US dairy industry is bovine respiratory disease (BRD) (USDA, 2007). This multifactorial disorder affecting the calf's respiratory tract has an even larger detrimental effect on the American beef industry, causing losses of approximately one billion dollars per year (Fulton et al, 2002; Griffin et al, 2010). Several pathogens, including viruses (Chase et al, 2008; Griffin et al, 2010) and bacteria (Shahriar et al, 2002; Thomson et al, 1975), comprise the BRD complex that causes pneumonia in calves and potentially death. Another relatively common disease in young dairy calves is infection of the middle ear (otitis media), which is highly correlated with pneumonia (Jensen 1983, Lima et al 2016). The upper respiratory tract (URT) seems to be an important link between these diseases (Jensen et al, 1983; Lima et al, 2016), since the nasopharynx communicates to the middle ear and resident microbes of the URT could be a source of bacteria for the lower respiratory tract (García-Rodríguez & Fresnadillo Martínez et al, 2002; Murphy et al, 2009). Furthermore, both pneumonia and otitis media share common risk factors, and the major bacteria

reportedly involved in the etiology of pneumonia are also associated with otitis media (Jensen et al, 1983). Our group recently described the assembly of the URT microbiome in newborn dairy calves, and determined that high total bacterial load present in the URT of 3-day-old calves, coupled with increased abundances of the bacterial genera *Mannheimia*, *Moraxella*, and *Mycoplasma*, is a key factor in the development of pneumonia in calves during the first 60 days of life (Lima et al, 2016).

The earliest potential source of microbial colonizers in newborn calves is the maternal microbiota, which can be acquired perinatally during passage through the birth canal and from contamination by maternal fecal microbes. Therefore, in the present study, we characterized the vaginal and fecal microbiotas of Holstein dairy cows within the last week of pregnancy and compared them to the fecal and URT microbiotas of their offspring to determine 1) the microbial taxa shared between dam and calf, and 2) the potential influence of the dam pre-partum vaginal and fecal microbiotas on calf URT and fecal microbial compositions at days 3, 14, and 35 of life. Additionally, we evaluated the potential influence of the dam's vaginal and fecal microbiota on the calf's respiratory tract and middle ear health during the pre-weaning period.

MATERIALS AND METHODS

Ethics statement

This study was conducted on a large commercial dairy farm located near Ithaca, New York. Animal Care and Use Procedures were according to the guidelines set out in Cornell University's Dairy Cattle Husbandry publication (no. 518). All experimental protocols using cattle were reviewed and approved by the Institutional Animal Care and Use Committee of Cornell University (Protocol numbers: 2013-0076 & 2011–0111).

Study design and study population

This study was a prospective observational cohort study in which 100 pregnant Holstein dairy cows from a single farm located near Ithaca, NY, were enrolled on a weekly basis. The farm was selected because of its longstanding relationship with the Ambulatory and Production Medicine Clinic at Cornell University. Vaginal and fecal samples were collected once at a single time pre-partum (273±3 days carried calf - DCC) for each cow under investigation and processed independently, as described in the next sections. Fecal and URT swabs from their respective progenies were longitudinally collected at days 3, 14 and 35 of life. After sample exclusion due to the following reasons—delivery of twins or a male calf, stillbirth, missed sample collection due to death of the calf before the end of the study and/or problems with downstream screening (e.g. failed DNA amplification or poor DNA sequence quality)—samples from 162 animals (81 cow–calf pairs) were subjected to further microbiota and statistical analyses.

Animals and facilities

Dry cows

Pregnant heifers and cows were housed together in free-stall barns with concrete stalls covered with rubber mattresses and bedded with dried manure solids, separated from the lactating cows. Pregnant animals were housed in two separate groups: the "far-off" group (where dry cows remained until two weeks before expected calving), and the "close-up" group (cows in their last 2 weeks before expected calving). Pregnant animals that were on stage 1 or 2 (stage 1 is define as: relaxation of the pelvic ligaments, dilatation of the cervix and distention of the teats; stage 2 is define as: delivery of the calf) of parturition were moved into the deep-bedded

maternity barn, with four identical group-pens (400 m^2 deep-bedded pens), to calve. Fresh cows were first milked within 8 hours of calving in a double 52-stall parallel milking parlor. Pre-partum cows were fed a diet with a high-fiber content and low energy density. The diet was formulated to meet or exceed the nutrient requirements for lactating Holstein cows weighing 650 kg and producing 45 kg of 3.5% fat corrected milk (NRC 2001).

Maternity

Calves were removed from the maternity pen and placed into a newborn pen bedded with dry sawdust and heated with heating lamps, right after parturition. Colostrum from multiparous (animals that experience more than one pregnancy) and primiparous (animals that experience only one pregnancy) cows was pooled and used in the study. All calves were fed approximately 4 L of colostrum at once by an esophageal feeder (Oral Calf Feeder Bag with Probe, Jorvet) within 4 hours of birth.

Calf barn

Twice daily, newborn calves were transferred from the newborn pen to the calf barn. The calf barn was a greenhouse type of barn with positive ventilation and divided into 18 identical group-pens. Group-pens had a total area of 70 m^2 and were bedded with straw on top of a thin layer of dry composted manure. Steel gates divided the group pens, and calves were allocated by birth order into each pen until the pen was completely full (a total of 25 calves per pen). All calves remained in the same pen from day 1 of life until fully weaned (approximately 65 days).

Calves were fed ad-libitum acidified non-saleable milk. The feeding system was fully automated. Briefly, the acidification was performed inside a sealed stainless-steel tank where the

non-saleable cold (5ºC) milk was mixed continuously with organic acid until pH 4.5 was reached. The acidified milk was kept for 72 hours inside the stainless-steel tank after the acidification process was finished. Then, the milk was directed to a smaller stainless-steel tank, which maintained the milk at a warm temperature and distributed it to the pens. To support the ad-libitum system, 6 nipples per pen were connected to the smaller tank and the acidified non-saleable milk was available from day 1 to day 55 of life, when a reduction of milk availability was initiated. All calves in this study were weaned by reducing the milk availability starting on day 55 until complete absence of acidified non-saleable milk at 65 days of life.

Sample collection

Study cows were identified and restrained in a headlock stanchion; the perineum and vulva were cleaned with a paper towel and disinfected with a 70% ethyl alcohol solution. The lips of the vulva were opened and a sterile swab (Puritan Medical Products, Guilford, ME) was applied to a single site at the midpoint of the vaginal cavity, swirled 6 times, and then withdrawn without touching surfaces outside the vagina. Each vaginal swab was immediately placed in a 2-ml microcentrifuge tube (VWR® International, Radnor, PA). Fecal samples were collected subsequently to the vaginal swab by using palpation gloves that were gently introduced in the rectum of the cows. Feces were then added to a sterile falcon tube (VWR® International, Radnor, PA). Both vaginal and fecal samples were kept on ice until they were transferred to the laboratory at the Cornell University College of Veterinary Medicine and stored at -80 °C.

Calf deep nasal pharyngeal swabs and fecal swabs were collected from each calf at days 3, 14, and 35 of life using a 20-cm sterile swab (Puritan Medical Products, Guilford, ME) covered by a thin sterile plastic sheath. Deep nasal pharyngeal swabs were collected as

previously described (Lima et al 2016). Briefly, prior to sampling, the calf was appropriately restrained and a nostril was cleaned using a paper towel. Subsequently, a plastic-covered swab was inserted into the nasal cavity, the plastic sheet was broken, and the swab was exposed to the URT mucosa. The tip of the swab was placed inside a sterile plastic tube and labeled. Next, under slight restraining, fecal swabs were obtained from each calf (a sterile cotton swab was inserted approximately 5 cm in the rectum). All fecal swabs were obtained at the same time as the respective URT swabs. URT and fecal swabs were frozen at -80 ºC until used for DNA extraction.

Case definition for calf pneumonia and otitis media

Diagnosis of pneumonia and otitis media was performed by experienced farm employees trained by Cornell University veterinarians (Ambulatory and Production Medicine Clinic) and confirmed by one of the veterinarians of the research team. Pneumonia was defined when two or more of the following clinical signs were detected in a calf: cough, rectal temperature >39.5 °C, respiratory rate >40 breaths/min, increased cranioventral lung sounds or wheezes. Otitis media was defined by observation of ear pain evidenced by head shaking, scratching or rubbing the ears, epiphora, ear droop, signs of facial nerve paralysis, with or without fever (rectal temperature >39.5 °C).

Calf health status was categorized as healthy (consisting of calves that did not develop pneumonia, otitis, and pneumonia-otitis combined during the pre-weaning period) or diseased (consisting of calves that developed pneumonia, otitis, or pneumonia-otitis combined during the pre-weaning period). Calves diagnosed with pneumonia and/or otitis were treated according to the standard farm protocol (Resflor Gold, Merck Animal Health).

DNA extraction

Isolation of DNA from cow vaginal swabs and calf URT and fecal swabs was performed by adding 1.5 ml of DNA/RNA-free ultra-pure water into the 2-ml microcentrifuge tubes (VWR® International, Radnor, PA) that contained URT, feces or vaginal swab samples. The tubes were then vortexed for 10 minutes by using a vortex adaptor (MO BIO Laboratory Inc., Carlsbad, CA) that holds 2-ml microcentrifuge tubes horizontally. The swabs were removed from the microcentrifuge tubes and the remaining liquid was centrifuged at room temperature for 5 minutes at 13,000 rpm. The supernatant was discarded and the DNA was extracted from the pellet using a PowerSoil DNA Isolation Kit (MO BIO Laboratory Inc., Carlsbad, CA) according to the manufacturer's recommendations.

Isolation of DNA from cow fecal samples was performed after homogenization of the feces content, and a total of 250 mg of feces solution was added directly to the PowerSoil beads tube, and DNA extraction was performed following the manufacturer's recommendations (MO BIO Laboratory Inc., Carlsbad, CA).

DNA concentration and purity were evaluated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 260 and 280 nm.

Amplicon sequencing and bioinformatics

16S rRNA amplification

Amplification of the V4 hypervariable region of the bacterial/archaeal 16S rRNA gene was performed For each sample evaluated by using the 515F and 806R primer set according to previously described methods and optimized for the Illumina MiSeq platform (Caporaso et al,

2012). PCR products were tagged with 280 different 12-bp error-correcting Golay barcodes (http://www.earthmicrobiome.org/) (Gilbert et al, 2010). PCRs were carried out in triplicate 25-µl reactions using 15-100 ng of template DNA, 1X EconoTaq Plus Green Master Mix (Lucigen, Middleton, WI) and 10 µM of each primer. Thermal cycling consisted of an initial denaturing step of 94°C for 3 min followed by 35 cycles of 94°C for 45 s, 50°C for 1 min and 72°C for 90 s, and a final elongation step of 72°C for 10 min. Replicate amplicons were pooled and visualized on 1.2% agarose gels stained with 0.5 mg/ml ethidium bromide, followed by purification using a Gel PCR DNA Fragment Extraction kit (IBI Scientific, Peosta, IA).

Amplicon concentration and purity were evaluated using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE) at wavelengths of 260 and 280 nm.

Sequence processing

Aliquots of fecal, vaginal and URT amplicon samples were standardized to the same concentration and pooled into 3 different sequencing runs according to individual barcode primers for the 16S rRNA gene. Final equimolar libraries were sequenced using the MiSeq reagent kit v2 (300 cycles) on the MiSeq platform (Illumina, Inc., San Diego, CA). The generated 16S rRNA gene sequences were demultiplexed using the Quantitative Insights Into Microbial Ecology 1 (QIIME) pipeline (Caporaso et al, 2010) version 1.9.1-dev. Sequences were filtered for quality using established guidelines (Bokulich et al, 2013). Sequences were binned into operational taxonomic units (OTUs) based on 97% identity using UCLUST (Edgar et al, 2010) against the Greengenes reference database (McDonald et al, 2012), May 2013 release. Low-abundance clusters were filtered and chimeric sequences were removed using USEARCH (Edgar et al, 2010). Phylogenetic trees were generated from the filtered alignment using FastTree

(Price et al 2010). Taxonomy was assigned using UCLUST (Edgar et al, 2010) consensus taxonomy assigner, against the Greengenes reference database (McDonald 2012) and only full-length, high-quality reads (-r = 0) were used for analysis. Additionally, we generated an OTU table using the MiSeq Reporter Metagenomics Workflow. The MiSeq Reporter classification is based on the Greengenes database, and the output of this workflow is a classification of reads at multiple taxonomic levels: kingdom, phylum, class, order, family, genus and species.

Diversity measurements and Unifrac clustering for Principal coordinates analysis (PCoA)

To calculate how many categories of taxa or lineages were detected in each individual sample and how those taxa or lineages were distributed within samples from dam fecal and vaginal samples, and from calf fecal and URT samples, Chao1 richness (number of species) (Chao et al, 1984) and Shannon diversity (species evenness) (Zand et al, 1976) indexes were generated using the QIIME 1 pipeline. To account for uneven sequencing depths across samples, all sample libraries were rarefied to an equal depth of 11,663 sequences before calculating the Shannon and Chao1 indexes.

To determine how taxa and or lineages were related within and between dam microbial communities and calf microbial communities, both phylogenetic unweighted and weighted UniFrac distances matrixes were generated in QIIME (Lozupone & Knight et al, 2005). UniFrac distances are based on the fraction of the branch length shared between communities within a phylogenetic tree constructed from the 16S rRNA gene sequences of all groups evaluated herein (dam vaginal and fecal microbiota, and calf fecal and URT microbiota). Unweighted UniFrac is a qualitative phylogenetic metric based on the presence or absence of bacteria, whereas weighted UniFrac is a quantitative phylogenetic metric incorporating bacterial relative abundance.

Principal coordinates were computed from the calculated UniFrac distance matrixes to compress dimensionality into three-dimensional principal coordinate analysis (PCoA) plots created by the "beta_diversity_through_plots.py" script in QIIME 1 and visualized by EMPeror (Vázquez-Baeza et al, 2013). To account for uneven sequencing depth across samples, all sample libraries were rarefied to an equal depth of 11,663 sequences before estimating the unweighted and weighted UniFrac distance matrixes.

Statistical analysis

Alpha diversity, Chao 1 richness and Shannon diversity indexes were calculated using QIIME 1. These diversity indexes were compared within and between dam and calf microbiotas using ANOVA in JMP Pro 12 (SAS Institute Inc.). Tukey-Kramer test was used to adjust for multiple comparisons.

The OTU data obtained from bioinformatics analysis were used to describe the relative abundances of bacterial phyla and genera within the dam and calf samples. Each value obtained indicates the percentage relative frequency of reads with 16S rRNA genes annotated to the indicated taxonomic level. The profiles of dam vaginal and fecal microbiota, and calf URT and fecal microbiota within each time point (3, 14 and 35 days of life) are described for the most prevalent phyla and genera using the tabulate function of JMP Pro 12 (SAS Institute Inc.). Relative abundance of bacterial phyla is presented in a stacked chart, and relative abundance of bacterial genera is presented in a pie chart.

Differences between microbial communities (beta diversity) based on phylogenetic information visualized on the PCoA plots were calculated with analysis of similarities (ANOSIM), a non-parametric statistical method (Clarke et al, 1993), in QIIME 1. ANOSIM with

999 permutations was used in this procedure to test for statistically significant differences between sample groups based on UniFrac distance matrixes. The values of the ANOSIM statistic R indicate the degree of separation across communities: R values closest to 1 suggest dissimilarity between groups, whereas R values closest to 0 suggest similarity between groups. The percentage of time that the actual R surpassed the permutation-derived R′ value is the p-value for the actual R statistic (Clarke et al, 1993; Clarke & Gorley et al, 2001). In the present study, a p-value ≤ 0.05 was considered significant.

The similarity of bacterial community composition was evaluated by using Venn diagrams (VennDiagram package under RStudio software version 0.99.903; RStudio, Inc) for graphical descriptions of unique and shared OTUs between and within dam and calf body sites. The average of shared OTU counts was analyzed by a general linear model (ANOVA) adjusted through the Tukey-Kramer multiple comparison correction in JMP Pro 12 (SAS Institute Inc.). After defining a core microbiome across dam and calf communities, we investigated which bacterial taxa mostly accounted for similarities between groups, based on rank abundances created for each group. Heatmaps were generated to graphically represent the relative distributions of the most common bacterial genera found in the core microbiomes of the dam and calf feces groups or the dam vaginal and calf URT groups.

To investigate whether the microbial interactions found between the dam vaginal and fecal microbial communities and the calf fecal and URT microbial communities play a role in controlling the fate of infection in calves, the relative abundances of the 40 most common bacterial genera detected in dam vaginal samples and dam fecal samples were used as covariates in a discriminant analysis model built in JMP Pro 12. Discriminant analysis predicts membership in a group or category based on observed values of several continuous variables (bacterial

relative abundance). Total canonical structure values, an output of the discriminant analysis that represents the correlation between the canonical variables and the covariates, were used to create the screening graphic.

A two-sided Welch's t-test followed by the Benjamini-Hochberg false discovery rate (FDR) was performed, using STAMP v. 2.1.3 (Parks et al, 2014), to examine the mean relative abundance and mean significant difference of each bacterial taxon selected by the screening model, as the major microbes differentiating the vaginal microbiota of dams whose calves developed pneumonia and/or otitis (disease) from dams whose calves did not develop pneumonia and/or otitis (healthy) during the pre-weaning period.

Based on those findings, we performed a Fisher's exact test in JMP Pro 12 (SAS Institute Inc.) to compare the incidence of disease (pneumonia and/or otitis) in calves born from cows with vaginal microbiota showing low *Mannheimia* abundance (relative abundance $\leq$ 1%), moderate *Mannheimia* abundance (1% < relative abundance $\leq$ 12.3%) or high *Mannheimia* abundance (relative abundance > 80%). These groups were created based on the normal quartiles of the *Mannheimia* relative abundance distribution detected in the dam vaginal samples. No sample presented *Mannheimia* relative abundance between 12.4% and 79.9%.

RESULTS

Sequencing results and descriptive statistics

The total post-quality-control number of sequences (sequences were filtered for size, quality, and for the presence of chimeras) used in the study was 53,564,884. The average coverage of sequences per sample was 83,434 (median = 69,134 sequences) with a SD of 51,255.

The incidences of pneumonia, otitis, and pneumonia-otitis combined in the studied calves during the pre-weaning period were 19.7% (n = 16), 34.5% (n = 28), and 6.2% (n = 5), respectively; 39.5% (n = 32) were considered healthy. The average age at first diagnosis was 23.6 (±1.4) days for pneumonia, 24.2 (±1.5) days for otitis, and 23.4 (±1.1) days for pneumonia-otitis combined.

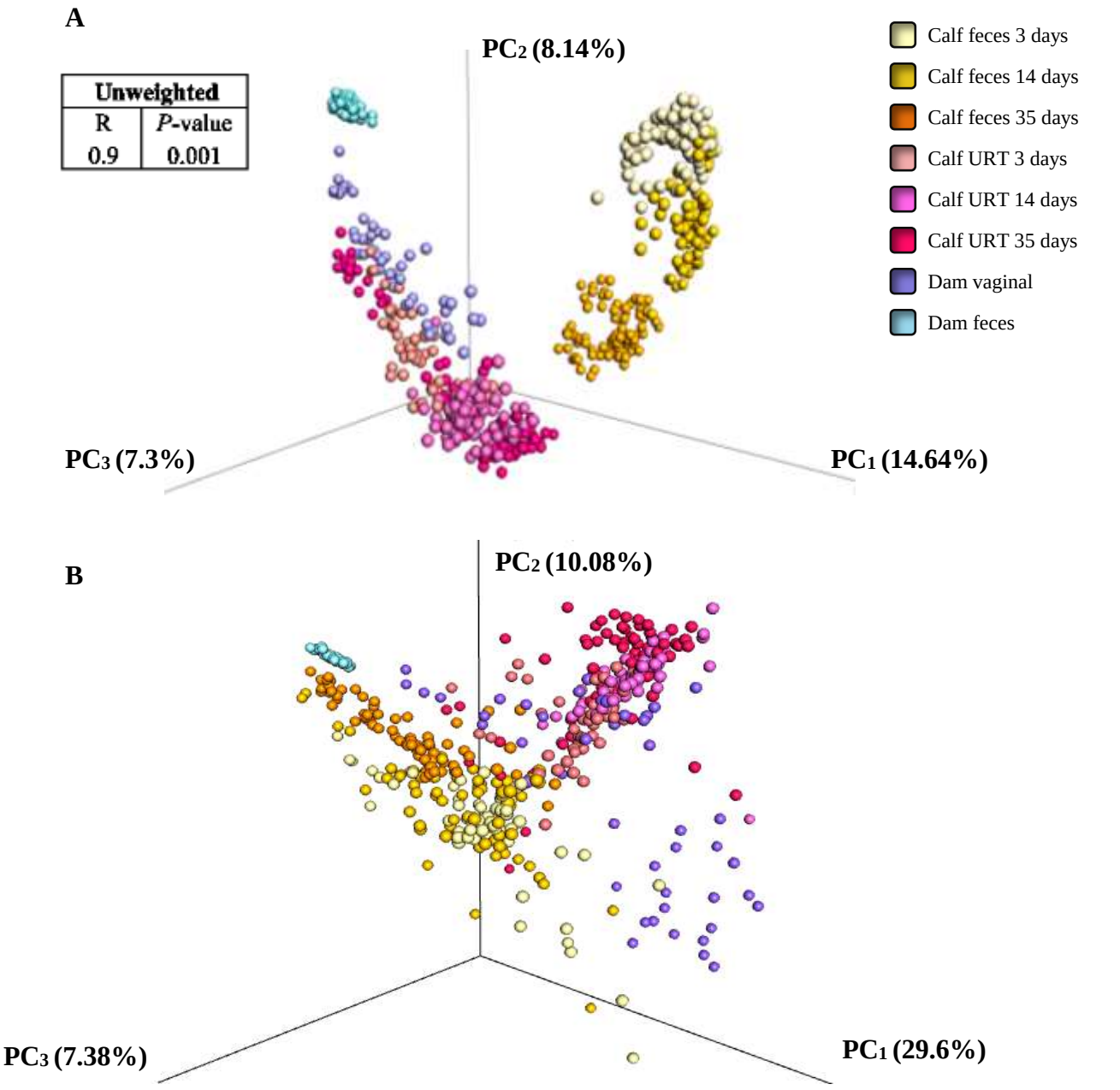
Characterization of microbial communities

The alpha diversity in dam and calf samples was determined by Chao 1 and Shannon indices and is depicted in Supplemental 5.1. The Chao 1 richness index (number of different species) was significantly different between the dam fecal and vaginal microbiotas (Supplemental 5.1A). No significant difference was found among the calf fecal microbiotas sampled at the different time point (days 3, 14 and 35 of life; Supplemental 5.1A). Additionally, comparisons between calf URT samples showed them to be significantly richer at day 3 versus day 14, but not at day 35. Calf feces, regardless of age, consistently had the lowest richness index of all the samples collected from both dams and calves. However, an effect of calf age was observed for the Shannon diversity (microbial evenness), which increased with neonatal age (Supplemental 5.1B). The calf URT samples showed significantly higher diversity on day 3 compared to day 14, followed by a further significant increase at day 35 compared to day 14 (Supplemental 5.1B).

The beta diversity analysis, which measures the level of similarities between samples as a function of microbial composition, was undertaken using UniFrac distance matrixes and allowed us to evaluate the potential impact of the dam microbiota on the early calf microbiota. When combining data from all sample types (dam vaginal swabs and feces; calf feces and URT swabs over time), the samples clustered primarily by the presence and absence of taxa (Figure 5.1A;

unweighted UniFrac R = 0.9, *P*-value = 0.001). This was not the case when OTU abundance was taken in consideration (Figure 5.1B; unweighted UniFrac R = 0.6, *P*-value = 0.001), which would suggest that the dissimilarities between the sample types were more accounted for by the presence or absence of OTUs than by their relative abundances.

Figure 5.1: Principal coordinate analysis (PCoA) of dam fecal and vaginal microbiota, and calf fecal and upper respiratory tract (URT) microbiota at 3, 14 and 35 days of life based on UniFrac distances matrices. The variance explained by each PCoA is given in parentheses. Each point corresponds to a microbial community colored according to each type of sample (dam fecal and vaginal; calf fecal and URT) and days of life (days 3, 14 and 35 of life). Comparisons of the UniFrac metrics are depicted in a) PCoA with unweighted UniFrac, performed on presence/absence and b) PCoA with weighted UniFrac, incorporating OTU abundances. The R statistic and *P*-value for differential clustering as assessed by the ANOSIM test, based on 999 permutations, are shown in the inset. The test statistic R can range from 1 to -1. An R value close to 1 suggests dissimilarity between groups, whereas an R value close to 0 suggests similarity between groups.



The dam vaginal microbiota was found to be highly similar to the calf URT microbiota regardless of calf age, which is clearly illustrated by the unweighted UniFrac PCoA plot (Table 5.1; R = 0. 37, *P*-value = 0.001). The differences (dissimilarity) in microbial composition between the dam vaginal and calf URT microbiotas account for only 37% of the variation between both microbiotas, indicating that the compositions of the dam vaginal and calf URT microbial communities overlapped by 63%. The level of similarity between the dam vaginal and the calf fecal microbiotas was much smaller (Table 5.1; unweighted UniFrac distance: R = 0. 72, *P*-value = 0.001 weighted UniFrac distance: R = 0. 91, *P*-value = 0.001), differing in composition by 72% (weighted UniFrac) and 91% (unweighted UniFrac). Nevertheless, when the bacterial taxa were plotted in the same PCoA space, based on the weighted average of the PCoA coordinates of all samples, where the weights are the relative abundances of the taxon in the samples, the dam fecal microbiota clustered closely to the calf fecal microbiota (Figure 5.1B and Table 5.1). Age was especially important, with a progression toward an adult-like state over the first 35 days of life (Figure 5.1B and Table 5.1).

Table 5.1: Analysis of similarities (ANOSIM) results for microbiota composition compared between sample types on weighted and unweighted UniFrac distance beta diversity. The test statistic R can range from 1 to -1. An R value close to 1 suggests dissimilarity between groups, whereas an R close to 0 suggests similarity between groups. Significance of the R statistic was determined by permuting group membership 999 times. When more than two groups were compared, in the case of calf feces and calf URT (upper respiratory tract), the samples were from all three age-groups (3, 14 and 35 days of life).

UniFrac distance	**ANOSIM**	
	R	*P*-value
Dam feces		
Unweighted		
Dam feces vs. Calf feces	0.92	0.001
Dam feces vs. Calf feces 35 days	0.99	0.001
Dam feces vs. Calf URT	0.72	0.001
Weighted		
Dam feces vs. Calf feces	0.61	0.001
Dam feces vs. Calf feces 35 days	0.39	0.001
Dam feces vs. Calf URT	0.93	0.001
Dam vaginal		
Unweighted		
Dam vaginal vs. Calf URT	0.37	0.001
Dam vaginal vs. Calf feces	0.72	0.001
Weighted		
Dam vaginal vs. Calf URT	0.36	0.001
Dam vaginal vs. Calf feces	0.91	0.001
Calf		
Unweighted		
Calf feces vs. Calf URT	0.65	0.001
Weighted		
Calf feces vs. Calf URT	0.43	0.001

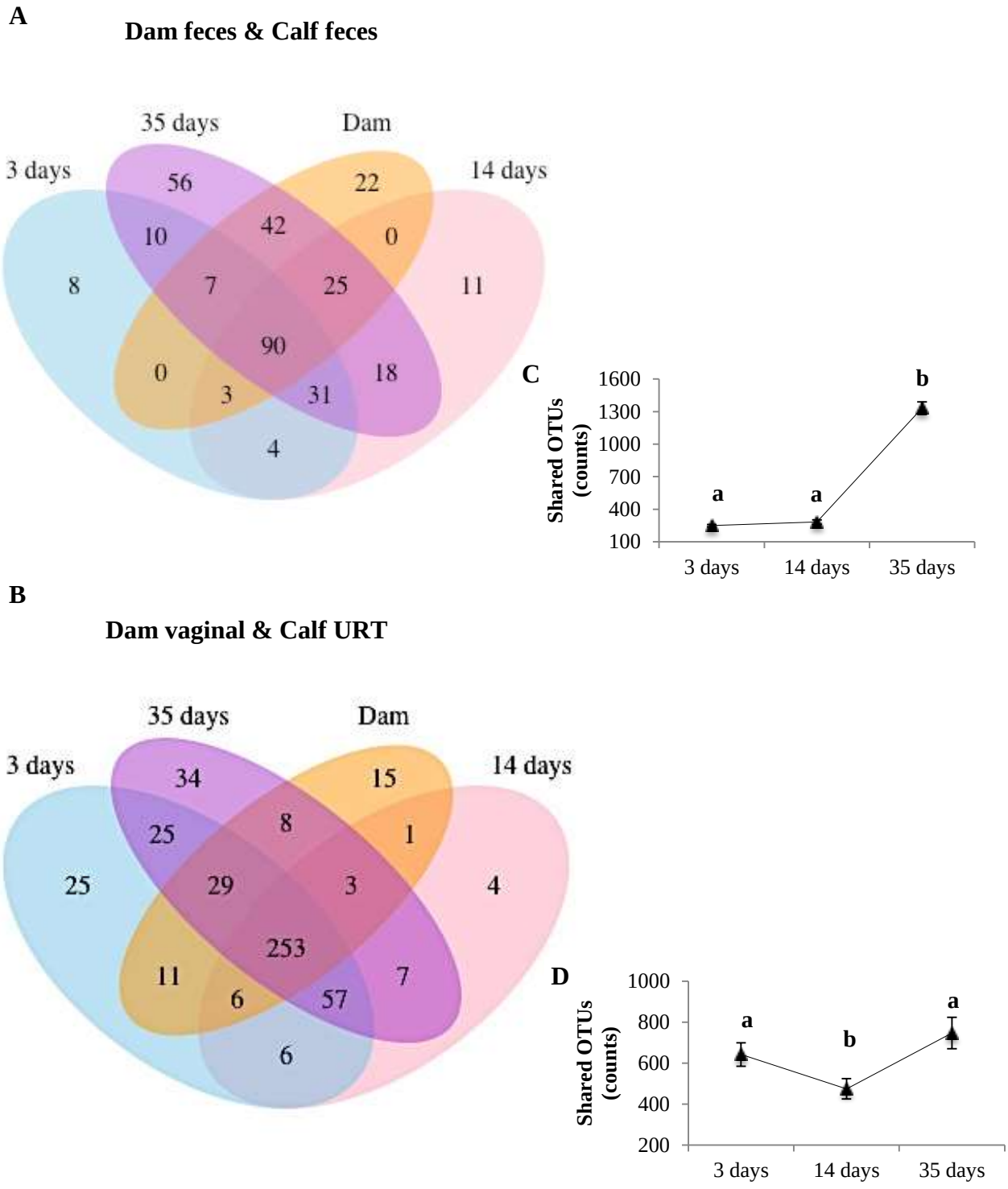
Analysis of the effect of parity (primiparous vs. multiparous) on the matrixes is depicted in Supplemental 5.2A and B. No dissimilarities were detected when the vaginal and fecal microbiotas of primiparous cows were compared to those of multiparous cows, regardless of the presence/absence of taxa (Supplemental 5.2B, unweighted UniFrac distance; R = 0.09, *P*-value = 0.005) or relative abundance of the taxa (Supplemental 5.2A, weighted UniFrac distance; R = 0.1, *P*-value = 0.003). A low degree of overlap was observed in bacterial community composition between the dam fecal and vaginal samples, which overlapped by 33% (Supplemental 5.3, unweighted UniFrac distance; R = 0.66, *P*-value = 0.001) and 34% (Supplemental 5.3, weighted UniFrac distance; R = 0.65, *P*-value = 0.001). Additionally, we found strong clustering of the adult gut represented by the fecal microbiota samples, suggesting that there is little variation in dam fecal microbiota (Supplemental 5.3).

The core microbiome framework within microbial communities

Quantification of the types and numbers of shared OTUs allowed us to evaluate the microbial relatedness among dams and calves. The Venn diagram showed that 90 OTUs were shared between the dam and calf fecal microbiotas, whereas 253 OTUs were shared by dam vaginal and calf URT microbiotas regardless of calf' days of life (Figure 5.2A and B, respectively). Sixty four, 65 and 60% of the detected OTUs were shared between maternal and calf fecal microbiota at days 3, 14 and 35 of life respectively, whereas 73, 76 and 87% were shared between the maternal vaginal microbiota and the calf URT microbiota at days 3, 14 and 35 of life, respectively. Following the same trend that we detected in the alpha diversity analysis, the number of OTUs (counts) shared by the dam and calf fecal microbiotas increased significantly over time (Figure 5.2C; *P*-value < 0.05), with the number at day 35 (post-birth)

being significantly higher compared to days 3 and 14. The number of shared taxa detected across the dam vaginal and calf URT microbiotas fell significantly between day 3 and day 14 (*P*-value < 0.05), but then increased significantly between days 14 and 35 (*P*-value < 0.05; Figure 5.2D).

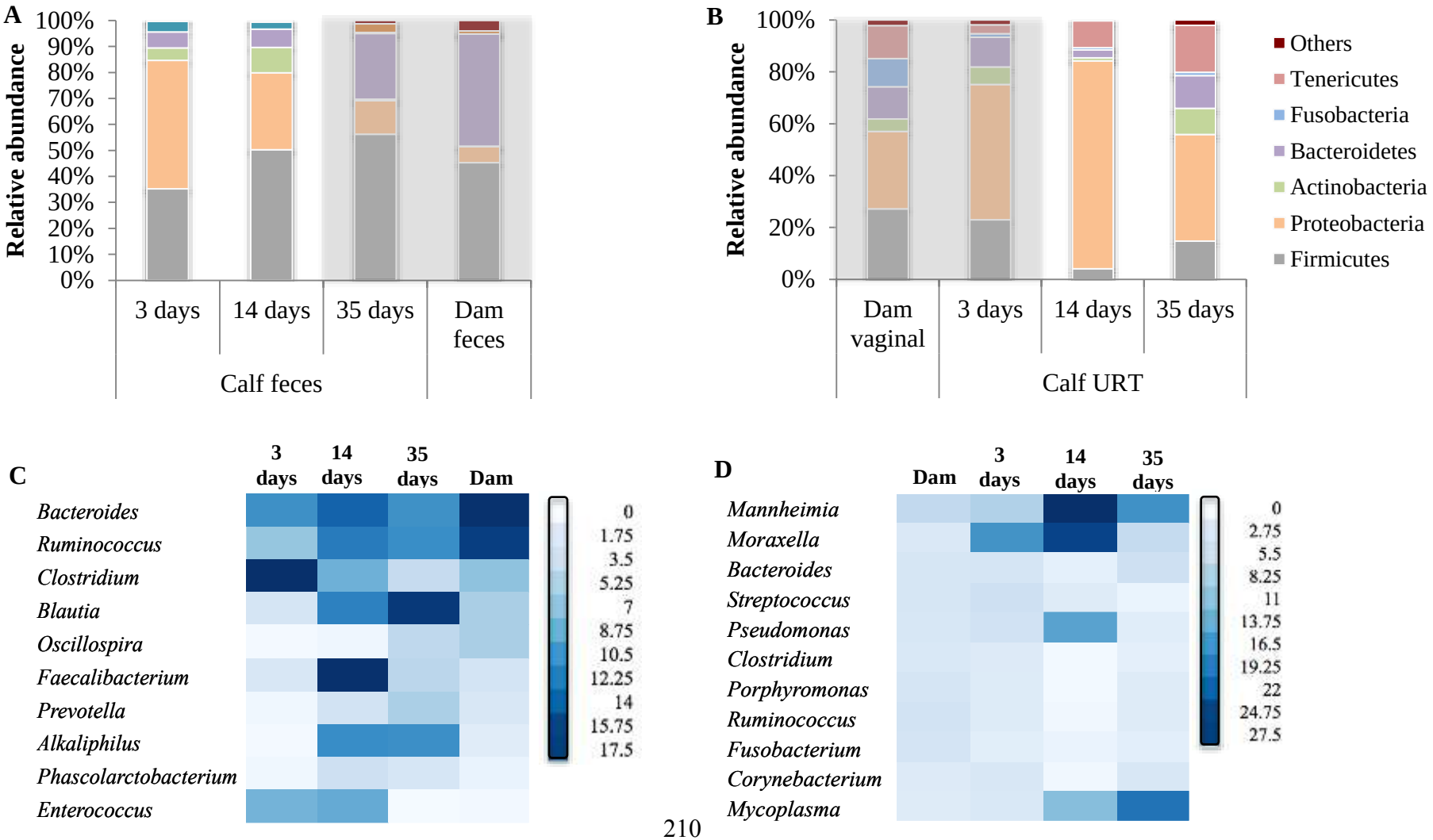
Figure 5.2: Venn diagram showing the numbers of unique and shared OTUs between dam feces and calf feces over time (3, 14 and 35 days of life) (A) as well as between dam vaginal and calf upper respiratory tract (URT) over time (B). Line graphs showing shared OTU counts between dam feces and calf feces (C) as well as between dam vaginal and calf URT over time (D). OTUs were defined at 97% sequence similarity. Error bars represent the 95% confidence interval. [a,b] different superscripts represent a significant difference ($P < 0.05$).



The dominant bacterial phyla identified in the dam vaginal and fecal microbiotas and calf fecal and URT microbiotas are illustrated in Figure 5.3A and B. At the phylum level, the fecal microbiotas of dams and 35-day-old calves were dominated by Firmicutes (45.3% ± 2.4 and 56.3% ± 18.7, respectively) and Bacteroidetes (43.1% ± 3.1 and 25.3% ± 15.5, respectively; Figure 3A). The most abundant phyla present in the dam vaginal microbiota and the calf URT microbiota at day 3 of life were: Proteobacteia (29.83% ± 22.3 and 52.1% ± 18.8, respectively); Firmicutes (27.2% ± 15.7 and 23% ± 9.6, respectively); Bacteroidetes (12.4% ± 9.1 and 11.4.3% ± 7.3, respectively; and Tenericutes (12.8% ± 12.0 and 3.4% ± 2.4, respectively) Figure 5.3B.

The core microbiota genera found in the dam and calf fecal microbiotas as well as in the dam vaginal and URT microbiotas are depicted in Figure 5.3C and D. *Bacteroidetes, Ruminococcus, Clostridium* and *Blautia* were the top four genera identified in the dam and calf fecal samples irrespective of calf age (Figure 5.3C). Notably, among the bacterial genera that were shared between the dam vaginal and calf nasopharyngeal samples, were those that are usually prevalent and considered important for the development of URT diseases in calves, such as *Mannheimia* and *Moraxella* (Figure 5.3D). *Mannheimia, Moraxella, Bacteroides, Streptococcus* and *Pseudomonas* were the top five genera identified within the most abundant bacterial genera in the dam vaginal and calf URT samples across all days of calf life (Figure 5.3D).

Figure 5.3: Microbial composition at the phylum level detected in dam and calf feces (A) as well as in dam vaginal and calf upper respiratory tract (URT) samples (B). The lower panels depict the 10 most common genera detected in the core microbiome of dam and calf feces (C) as well as dam vaginal and calf URT samples (D). The core microbiome was categorized as being the bacterial genera detected in all samples of dam and calf feces (C) or dam vaginal and calf URT (D). Each square in the heatmap represents the abundance level of a single category. Small relative abundance values are white, progressing to higher values as dark blue.

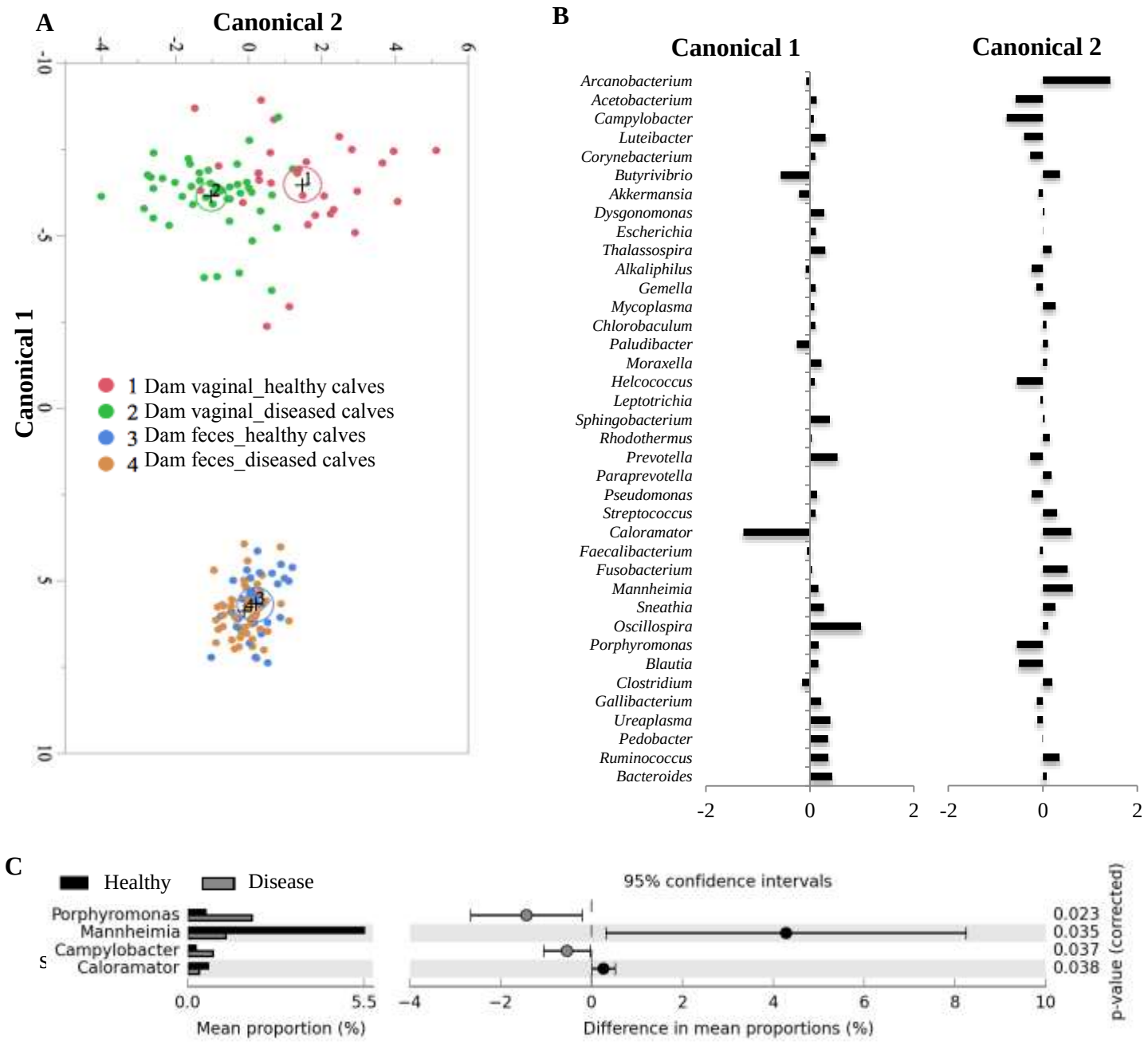


Detailed descriptions of the most common bacterial phyla and genera detected in dam feces and dam vaginal samples are given in Supplemental Figure 5.4A, B and C. The most common genera detected in fecal samples of calves at days 3 (Supplemental 5.5A), 14 (Supplemental 5.5B) and 35 (Supplemental 5.5C), as well as in URT samples at days 3 (Supplemental 5.5D), 14 (Supplemental 5.5E) and 35 (Supplemental 5.5F) are described in detail in Supplemental 5.5.

The effect of dam microbiota on calf health

We evaluated whether there existed a unique microbiota of maternal fecal and vaginal bacterial communities shared by all dairy cows in this study that was associated with health of the respiratory tract and/or middle ear of calves (healthy vs. pneumonia and/or otitis) during the pre-weaning period. A discriminant analysis based on dam sample type and calf health is illustrated in Figure 5.4A and was used to identify the most important maternal bacteria involved in the development of disease (pneumonia and/or otitis) in the neonatal calves. The discriminant function is used to calculate canonical scores that describe the separation between groups. The differences between dam fecal and vaginal microbiota as a function of offspring health status are depicted in the canonical scores 1 and 2 (Figure 5.4B). The results of a more detailed analysis of the differences in dam vaginal bacterial genera between calves that developed disease (pneumonia and/or otitis) and those that did not (healthy calves) are shown in Figure 5.4C. The genera *Porphyromonas*, *Campylobacter*, and *Caloramator* were relatively more abundant in the vaginal microbiota of dams whose progeny did not develop URT disease, and *Mannheimia* was relatively more abundant in the vaginal microbiota of dams whose progeny developed URT disease (Figure 5.4C).

Figure 5.4: Discriminant analysis of dam feces and vaginal microbiotas according to calf upper respiratory tract (URT) health status (healthy calves, or calves that developed pneumonia, otitis, or both diseases combined during the pre-weaning period, termed 'diseased calves'). The 40 most common shared bacterial genera were added to the discriminant procedure. The ellipses represent the 95% confidence region that contains the true mean of the group, and a plus symbol indicates the center (centroid) of each group (A). Differences in the dam microbial profiles for each health group and sample type detected in the discriminant analysis are illustrated by Canonicals 1 and 2 (B). Detailed bacterial genera differences between dam vaginal microbial groups based on calf health status are illustrated by the x-axis values (C). Welch's test with Benjamin-Hochberg FDR correction was applied on these datasets (C). The results were filtered using a *P*-value of 0.05 and an effective size of 0.5 threshold in STAMP (C).



we categorized the vaginal microbiota into 3 groups: low *Mannheimia* relative abundance (range = 0.98 – 0.01%, N = 39), moderate *Mannheimia* relative abundance (range = 1.05 – 12.13%, N = 32), and high *Mannheimia* relative abundance (range 80.15% – 87.67%, N = 4; Table 2). The incidence of disease (pneumonia and/or otitis) in calves born from cows in each category is depicted in Table 5.2.

Table 5.2: *Mannheimia* relative abundance in the vagina of dams and its association with calf health status (healthy vs. animals that developed pneumonia and/or otitis, described as 'disease') during the pre-weaning period. Dam vaginal microbiota according to the prevalence (relative abundance) of *Mannheimia* is categorized as low: *Mannheimia* relative abundance ≤ 1%; moderate: 1 < *Mannheimia* relative abundance ≤ 12.13%; and high: *Mannheimia* relative abundance > 80%. N represents the number of calves in each category described, and the *P*-value refers to the difference between the *Mannheimia* relative abundance groups.

***Mannheimia* prevalence in the cow vagina**	**Pneumonia and/or otitis incidence**
Mannheimia RA[1] ≤ 1%	68.3% (28)
1 < *Mannheimia* RA[1] ≤ 12%	56.7% (17)
Mannheimia RA[1] > 80%	0
Comparison between groups: low, moderate and high vaginal *Mannheimia* abundance	***P*-value**
Mannheimia RA[1] ≤ 1% vs. 1 < *Mannheimia* RA[1] ≤ 12%	0.33
Mannheimia RA[1] ≤ 1% vs. *Mannheimia* RA[1] > 80%	0.01
1 < *Mannheimia* RA[1] ≤ 12% vs. *Mannheimia* RA[1] > 80%	0.10

[1]RA: Relative abundance

DISCUSSION

To our knowledge, this study is the first to use a culture-independent technique to evaluate the influence of the maternal vaginal and fecal microbiotas on the microbiota of neonatal calves. Our results indicate that the vaginal microbiota of the dam is a significant source

of the bacteria that colonize the URT of the newborn calf (at day 3 of life), and its influence seems to persist, at least until day 35 of life. A total of 253 taxa were shared between the dam vaginal and calf URT samples, and *Mannheimia* and *Moraxella,* which are considered important bacteria for the development of respiratory tract and middle ear diseases in calves (Lima et al 2016), were the most prevalent bacterial genera in this shared bacterial core. This association suggests that the health status of the calf's early respiratory tract and middle ear might be affected by mother-to-offspring transmission of bacteria. Furthermore, the fecal microbiota of the newborn calf differed from the fecal microbiota of the dam; however, the calf fecal microbial community changed relatively quickly, such that by day 35 of life it resembled the microbiota observed in the mature bovine gut (as represented by feces).

In this study, we used the UniFrac metric to determine biologically meaningful patterns of similarity between the dam vaginal and fecal microbiotas in relationship to the calf URT and fecal microbiotas and to better understand the origin and formation of the calf gut and URT microbial community. We observed a clear increase in the diversity of the calf gut microbiota over the 5-week study period. This finding was expected based on what was already known about the physiology and structure of the gut of young ruminants and the direct impact these have on gut microbial composition (Van Soest, 1994). In addition, changes in the diet, characterized by increased consumption of solid food and reduced ingestion of milk, is typical during this initial phase. Fecal bacterial composition of pre-weaning calves from 2 weeks of life until 1 year old was recently evaluated by Dill-McFarland et al (2017) and an increase in alpha-diversity (Shannon index) and decrease in beta-diversity as the calves aged was detected, which corroborate with our findings (Dill-McFarland et al, 2017). In the present study, at day 3 of life, the calf fecal microbiota was dominated by the phyla Proteobacteria and Firmicutes; however, at

day 35 of life, we observed a reduction of Proteobacteria and an increase in the relative abundance of Firmicutes and Bacteroidetes, resembling the microbiota present in the dam fecal samples. The genera *Bacteroides, Ruminococcus, Clostridium* and *Blautia* were also found to be part of the 15 most abundant genera in the dam fecal microbiota and in the calf fecal microbiota at 35 days of life.

In our analysis of similarity, the composition of the bovine vaginal and fecal microbiomes collected within a week prior to parturition did not strongly cluster. Conventionally, the strong similarity between the microbiotas of the bovine reproductive and digestive tracts has been attributed to several factors, including their anatomical proximity caudally, perineal hygiene, presence of bacteria in the vagina of healthy dairy cows that are directly or indirectly associated with the cow digestive tract, and the fact that many reproductive tract diseases are caused by microorganisms found in the feces of these animals (Dohmen et al, 2000; Gautam and Nakao et al, 2009; Schuenemann et al, 2013). However, this association was not evident in our study. Furthermore, calf fecal samples and calf URT samples showed dissimilar bacterial compositions. Nevertheless, when bacterial abundance was included in the beta diversity model, it was possible to visualize some similarities between the calf fecal and URT microbial communities. These findings partially disagree with those of Dominguez-Bello et al. (2010), who reported that human newborn nasopharyngeal and meconium microbial communities are essentially undifferentiated (Dominguez-Bello et al, 2010). The disparate findings between the two studies might be attributed to physiological differences between humans and cattle animal species, and the former evidence that meconium harbored a unique microbial community that was distinct from fecal samples in the first month of life (Moles et al, 2013).

In contrast to the evident temporal evolution observed in the calf gut microbiome, we saw no dramatic differentiation of the calf URT microbial community, even when OTU abundance was taken into consideration using the weighted UniFrac metric. Furthermore, the calf URT at days 3 and 35 of life harbored a more diversified microbial community (composed mainly of Proteobacteria, Firmicutes, Bacteroidetes and Actinobacteria) than the microbiota observed at day 14, which was dominated by members of the Proteobacteria phylum, representing 80% of the total microbial community. This increased relative abundance of Proteobacteria was mainly driven by the increased relative abundance of the *Mannheimia* and *Moraxella* genera. Thus, the convention that respiratory-tract communities dominated by few organisms also exhibit low evenness (Laurenzi et al, 1961) is supported by our findings of the increased relative abundance of *Mannheimia* and *Moraxella* and the lower microbial diversity index detected in the URT samples at day 14. The shift of the microbial profile detected at day 14 followed by a restoration at day 35 is likely due to what has been called the "window of susceptibility" encountered by young calves before weaning (Chase et al, 2008). This period characterized by a lower blood titer of maternal antibodies in conjunction with a still-immature calf immune system (Chase et al, 2008), conditions which might favor the overgrowth of particular bacterial species, resulting in the reduced bacterial diversity we observed at day 14 of life.

In the present study, we demonstrated that the dam vaginal microbiome is similar to the calf URT microbiome, regardless of whether bacterial abundance or presence/absence is taken into account. Our data indicate that the bovine birth canal is a source of microbial colonizers of the neonatal calf, and the similarity between the dam vaginal samples and the calf URT samples at all three days of neonatal life that we evaluated is evidence that the effect of the dam microbiota on the calf URT can persist. Our results are in accordance with recent literature

showing that substantial bacterial colonization of the human newborn occurs at birth upon exposure to vaginal and skin microbiota (Biasucci et al, 2010; Dominguez-Bello et al, 2010). In the study by Dominquez-Bello et al. (2010), babies born by vaginal delivery harbored a bacterial community profile, including the URT microbiome, similar to the vaginal community of their mothers. In the present study, 253 OTUs were shared between the dam vaginal and calf URT microbiomes regardless of the calf day of life sampled. Additionally, the dam vaginal microbial profile was mostly composed of Firmicutes, Proteobacteria, Bacteroidetes, and Tenericutes phyla, which is comparable to the microbiota profile detected in the URT of the neonatal calf, especially at days 3 and 35 of life. Similarities were also detected at the genera level, in which bacteria such as *Bacteroidetes, Mannheimia, Moraxella, Streptococcus* and *Pseudomonas* were found to be highly abundant in both the dam vaginal and calf URT samples. The microbial composition of the cow pre-partum vagina as detected in the present study is similar to that recently described by Bicalho et al. (2017) in their study of the dynamics of the vaginal microbiome of dairy cows before and after parturition (Bicalho et al, 2017).

We also previously investigated the microbial evolution of the dairy calf URT and its association with respiratory disease and otitis media, and the most common pathogens associated with pneumonia and otitis media were already present in the URT of newborn calves at 3 days of life (Lima et al, 2016). Thus, in the present study, we investigated whether the dam vaginal microbiota would be a potential risk factor affecting the health of the calf respiratory tract and middle ear during the same period evaluated in our former study. Our discriminant analysis showed that the vaginal microbiota of dams whose calves went on to develop respiratory disease and/or otitis media tended to be separate from the microbiota of dams whose calves stayed healthy. Furthermore, and of great significance, bacterial genera such as *Mannheimia,*

Mycoplasma, Streptococcus, and *Moraxella* were detected in the vaginal microbiota of dams, and members of these genera are important contributors to bovine respiratory and/or otitis media disease in calves (Catry et al , 2007; Griffin et al, 2010; Lima et al, 2016; Yeruham et al, 1999).

Surprisingly, and worth noting in the present study, the genus *Mannheimia* was found to be more prevalent in the vaginal microbiota of dams whose calves did not develop pneumonia and/or otitis media disease compared to the microbiota of dams whose calves did develop disease. It appears that the pre-partum higher abundance of *Mannheimia* in the vagina of dairy cows might have a protective effect on the health of the respiratory tract and middle ear of their progeny. Given this interesting finding, we may suggest that antibodies generated in response to microbial colonization of the dam's vagina shape the composition of the calf microbiota in ways beneficial to the URT health of calves. The mucosal immune system has been shown to be a potential route for effective and long-lasting vaccination against respiratory pathogens, because it benefits from the poor interference of mostly secretory antibodies acquired from colostrum (Chase et al, 2008). This may reflect a natural immunization process involving nasal delivery of dam vaginal microbes to neonatal calves. Furthermore, Lee et al. (2015) showed that a robust population of dendritic cells patrols the nasal cavity in mice and humans and is poised to respond to antigen presence (Lee et al, 2015). Thus, the protective effect observed in our study by the dam vaginal microbiota might be potentially explained by an induced antigen-specific immunity across calf nasal mucosa site. In fact, when we categorized the dam vaginal samples in groups of low-, moderate- and high-abundance *Mannheimia,* all progeny from the dams categorized in the high-abundance *Mannheimia* group (relative abundance > 80%) did not develop pneumonia and/or otitis during the pre-weaning period. However, only four cows presented a vaginal habitat dominated by this bacterial genus, suggesting that the protective effect might require a high titer

of *Mannheimia* in the vagina. Further studies are needed to confirm the protective effect of the dam vaginal microbiota on the health of the calf respiratory tract and middle ear during the pre-weaning period.

The role of immune recognition pathways in antibody-mediated immunoselection of the microbiota has been previously discussed in detail (Kubinak and Round et al, 2016). Kubinak and Round et al. (2016) discussed for instance how the maternal antibodies that are passed in milk to her progeny might influence gut microbiota composition (Kubinak and Round et al, 2016). The adaptive immune system does not only influence microbial fitness promoting microbial diversity and mediating symbiosis, but also function to reduce the expression of potentially harmful epitopes of commensals and prevent lethal dissemination by controlling mucosa associated communities (Kubinak and Round, 2016). Because in our study calves were fed with pooled colostrum collected from multiple fresh cows, following a standardized milk collection protocol the definition and characterization of the initial stages of calf URT microbiota symbiosis has to be determined by other pathways other than immune passive transfer. Recent studies have mounted a foundation of evidence for maternal microbial transmission in mammals that goes beyond the conventional ascension of vaginal microbes, to arguing against the prevailing dogma of a sterile placenta and colostrum, to further exploitation that microbial translocation is facilitated by dendritic cells (Jimenez et al., 2005, Gosalbes et al 2013, Aagaard et al., 2014). Therefore, despite the aforementioned evidence, it is still unclear as to what extent the microbiota found in the vagina of cows influences the microbiota composition in their offspring URT.

CONCLUSION

Our results indicate that the composition of the dam vaginal and calf URT microbial communities overlap by 63%. The dam vaginal microbiota appears to be vertically transmitted to the URT of the newborn calf (3 days of life) and to persist, at least until 35 days of life. The genera *Mannheimia*, *Streptococcus* and *Moraxella*, which are the most common bacterial genera associated with pneumonia and otitis media in calves, were among the most abundant genera in the dam vaginal microbiota. Additionally, the genus *Mannheimia* was relatively more abundant in the vaginal microbiota of dams whose progeny did develop URT disease. Dam and calf fecal microbiotas differed in their composition by 72% (weighted UniFrac) and 91% (unweighted UniFrac). Nevertheless, when relative abundance was taken into account, the dam fecal microbiota clustered closely to the calf fecal microbiota. Age was especially important, with progression toward an adult-like state over the first 35 days of life. Together, these results provide an unprecedented understanding of the impact of the dam microbiome on the initial microbial colonization of the neonatal calf, and emphasize the need for new studies to understand the potential protective effect of the dam vaginal microbiota on calf respiratory tract and middle ear health.

REFERENCES

Aagaard KM, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. (2014). The placenta harbors a unique microbiome. *Sci Transl Med* 6: 237.

Addis MF, Tanca A, Uzzau S, Oikonomou G, Bicalho RC, Moroni P (2016). The bovine milk microbiota: insights and perspectives from -omics studies. *Mol Biosyst*.

Biasucci G, Rubini M, Riboni S, Morelli L, Bessi E, Retetangos C (2010). Mode of delivery affects the bacterial community in the newborn gut. *Early Hum Dev* 86 Suppl 1: 13-15.

Bicalho ML, Santin T, Rodrigues MX, Marques CE, Lima SF, Bicalho RC (2017). Dynamics of the microbiota found in the vaginas of dairy cows during the transition period: Associations with uterine diseases and reproductive outcome. *J Dairy Sci*.

Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R *et al* (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10: 57-59.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7: 335-336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N *et al* (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6: 1621-1624.

Catry B, Boyen F, Baele M, Dewulf J, de Kruif A, Vaneechoutte M *et al* (2007). Recovery of *Moraxella ovis* from the bovine respiratory tract and differentiation of *Moraxella* species by tDNA-intergenic spacer PCR. *Vet Microbiol* 120: 375-380.

Chao A (1984). Nonparametric estimation of the number of classes in a population. *Scand J Stat*: 265-270.

Chase CC, Hurley DJ, Reber AJ (2008). Neonatal immune development in the calf and its impact on vaccine response. *Vet Clin North Am Food Anim Pract* 24: 87-104.

Cho I, Yamanishi S, Cox L, Methé BA, Zavadil J, Li K *et al* (2012). Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* 488: 621-626.

Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB *et al* (2012). Gut immune maturation depends on colonization with a host-specific microbiota. *Cell* 149: 1578-1593.

Clarke KR (1993). Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol*: 117-143.

Clarke KR, Gorley RN (2001). Primer v5 (Plymouth routines in multivariate ecological research): User manual/tutorial. Primer-E Ltd., Plymouth, UK.

Cox LM, Yamanishi S, Sohn J, Alekseyenko AV, Leung JM, Cho I *et al* (2014). Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell* 158: 705-721.

Dill-McFarland KA, Breaker JD, Suen G (2017). Microbial succession in the gastrointestinal tract of dairy cows from 2 weeks to first lactation. *Sci Rep* 7: 40864.

Dohmen MJ, Joop K, Sturk A, Bols PE, Lohuis JA (2000). Relationship between intra-uterine bacterial contamination, endotoxin levels and the development of endometritis in postpartum cows with dystocia or retained placenta. *Theriogenology* 54: 1019-1032.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N *et al* (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci USA* 107: 11971-11975.

Edgar RC (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-2461.

Fulton RW, Cook, B. J., Step, D. L., Confer, A. W., Saliki, J. T., Payton, M. E., Burge, L. J., Welsh, R. D., Blood, K. S. (2002). Evaluation of health status of calves and the impact on feedlot performance: assessment of a retained ownership program for postweaning calves. *Can J Vet Res* 66: 173-180.

García-Rodríguez JA, Fresnadillo Martínez MJ (2002). Dynamics of nasopharyngeal colonization by potential respiratory pathogens. *J Antimicrob Chemother* 50 Suppl S2: 59-73.

Gautam G, Nakao T (2009). Prevalence of urovagina and its effects on reproductive performance in Holstein cows. *Theriogenology* 71: 1451-1461.

Gilbert JA, Meyer F, Antonopoulos D, Balaji P, Brown CT, Brown CT, Desai N *et al* (2010). Meeting report: the terabase metagenomics workshop and the vision of an Earth microbiome project. *Stand Genomic Sci* 3: 243-248.

Gosalbes MJ, Llop S, Vallès Y, Moya A, Ballester F, Francino MP (2013). Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. *Clin Exp Allergy* 43: 198-211.

Griffin D, Chengappa MM, Kuszak J, McVey DS (2010). Bacterial pathogens of the bovine respiratory disease complex. *Vet Clin North Am Food Anim Pract* 26: 381-394.

Holman DB, McAllister TA, Topp E, Wright AD, Alexander TW (2015). The nasopharyngeal microbiota of feedlot cattle that develop bovine respiratory disease. *Vet Microbiol* 180: 90-95.

Jami E, Israel A, Kotser A, Mizrahi I (2013). Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J* 7: 1069-1079.

Jensen R, Maki LR, Lauerman LH, Raths WR, Swift BL, Flack DE, Hoff RL *et al* (1983). Cause and pathogenesis of middle ear infection in young feedlot cattle. *J Am Vet Med Assoc* 182: 967-972.

Jiménez E M L, Marín R L, Martín R, Odriozola J M, Olivares M, Xaus J. *et al* (2008). Is meconium from healthy newborns actually sterile? *Res Microbiol* 159: 187-193.

Kelly D, King T, Aminov R (2007). Importance of microbial colonization of the gut in early life to the development of immunity. *Mutat Res* 622: 58-69.

Kubinak, J L, Round JL (2016). Do antibodies select a healthy microbiota? *Nat Rev Immunol* 16: 767-774.

Laurenzi GA, Potter RT, Kass EH (1961). Bacteriologic flora of the lower respiratory tract. *N Engl J Med* 265: 1273-1278.

Lee H, Ruane D, Law K, Ho Y, Garg A, Rahman A *et al* (2015). Phenotype and function of nasal dendritic cells. *Mucosal Immunol* 8: 1083-1098.

Lif Holgerson P, Harnevik L, Hernell O, Tanner AC, Johansson I (2011). Mode of birth delivery affects oral microbiota in infants. *J Dent Res* 90: 1183-1188.

Lima FS, Oikonomou G, Lima SF, Bicalho ML, Ganda EK, Filho JC *et al* (2015). Prepartum and postpartum rumen fluid microbiomes: characterization and correlation with production traits in dairy cows. *Appl Environ Microbiol* 81: 1327-1337.

Lima SF, Teixeira AG, Higgins CH, Lima FS, Bicalho RC (2016). The upper respiratory tract microbiome and its potential role in bovine respiratory disease and otitis media. *Sci Rep* 6: 29050.

Lozupone C, Knight R (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228-8235.

Machado VS, Oikonomou G, Bicalho ML, Knauer WA, Gilbert R, Bicalho RC (2012). Investigation of postpartum dairy cows' uterine microbial diversity using metagenomic pyrosequencing of the 16S rRNA gene. *Vet Microbiol* 159: 460-469.

McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A *et al* (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6: 610-618.

Meale SJ, Li S, Azevedo P, Derakhshani H, Plaizier JC, Khafipour E *et al* (2016). Development of Ruminal and Fecal Microbiomes Are Affected by Weaning But Not Weaning Strategy in Dairy Calves. *Front Microbiol* 7: 582.

Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W *et al* (2013). Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. *PLoS One* 8: e66986.

Morein B, Abusugra I, Blomqvist G (2002). Immunity in neonates. *Vet Immunol Immunopathol* 87: 207-213.

Murphy TF, Bakaletz LO, Smeesters PR (2009). Microbial interactions in the respiratory tract. *Pediatr Infect Dis J* 28: S121-126.

NRC (2001). Nutrient Requirements of Dairy Cattle: Natl. Acad. Press, Washington, DC.

Oikonomou G, Bicalho ML, Meira E, Rossi RE, Foditsch C, Machado VS *et al* (2014). Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. *PLoS One* 9: e85904.

Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30: 3123-3124.

Price MN, Dehal PS, Arkin AP (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5: e9490.

Renz-Polster H, David MR, Buist AS, Vollmer WM, O'Connor EA, Frazier EA *et al* (2005). Caesarean section delivery and the risk of allergic disorders in childhood. *Clin Exp Allergy* 35: 1466-1472.

Ridge JP, Fuchs EJ, Matzinger P (1996). Neonatal tolerance revisited: turning on newborn T cells with dendritic cells. *Science* 271: 1723-1726.

Salam MT, Margolis HG, McConnell R, McGregor JA, Avol EL, Gilliland FD (2006). Mode of delivery is associated with asthma and allergy occurrences in children. *Ann Epidemiol* 16: 341-346.

Santos TM, Bicalho RC (2012). Diversity and succession of bacterial communities in the uterine fluid of postpartum metritic, endometritic and healthy dairy cows. *PLoS One* 7: e53048.

Schuenemann GM, Bas S, Gordon E, Workman JD (2013). Dairy calving management: description and assessment of a training program for dairy personnel. *J Dairy Sci* 96: 2671-2680.

Shahriar FM, Clark EG, Janzen E, West K, Wobeser G (2002). Coinfection with bovine viral diarrhea virus and *Mycoplasma bovis* in feedlot cattle with chronic pneumonia. *Can Vet J* 43: 863-868.

Thomson RG, Chander S, Savan M, Fox ML (1975). Investigation of factors of probable significance in the pathogenesis of pneumonic pasteurellosis in cattle. *Can J Comp Med* 39: 194-207.

USDA (2007). Heifer Calf Health and Management Practices on U.S. In: Agriculture (ed). USDA:APHIS:VS:CEAH NRRC Building B, M.S. 2E7 2150 Centre Avenue Fort Collins, CO 80526-811.

Van Soest P (1994). *Nutritional ecology of the ruminant.*, 2 edn. Cornell University Press: New York. USA.

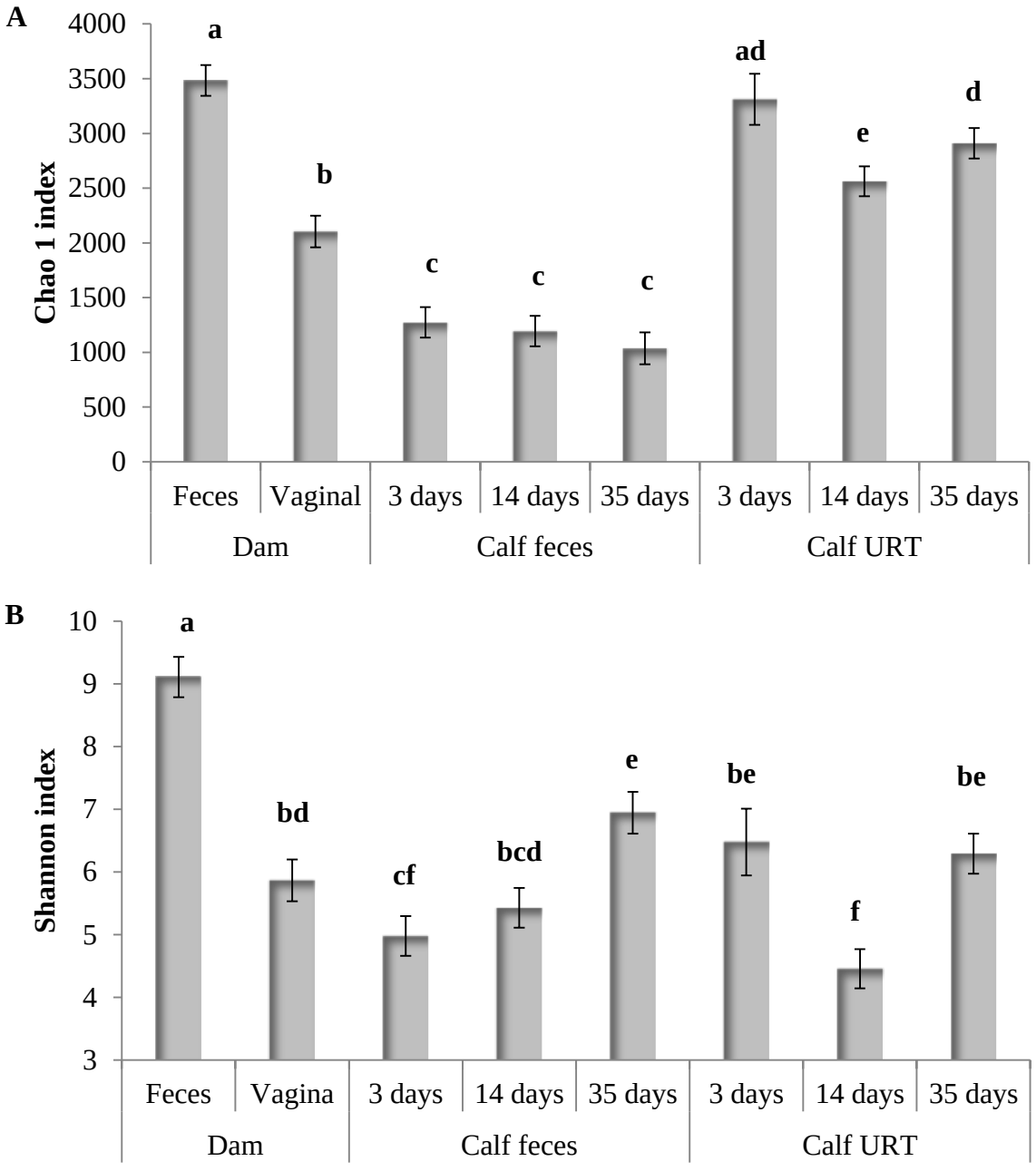
Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R (2013). EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* 2: 16.

Yeruham I, Elad D, Liberboim M (1999). Clinical and microbiological study of an otitis media outbreak in calves in a dairy herd. *Zentralbl Veterinarmed B* 46: 145-150.

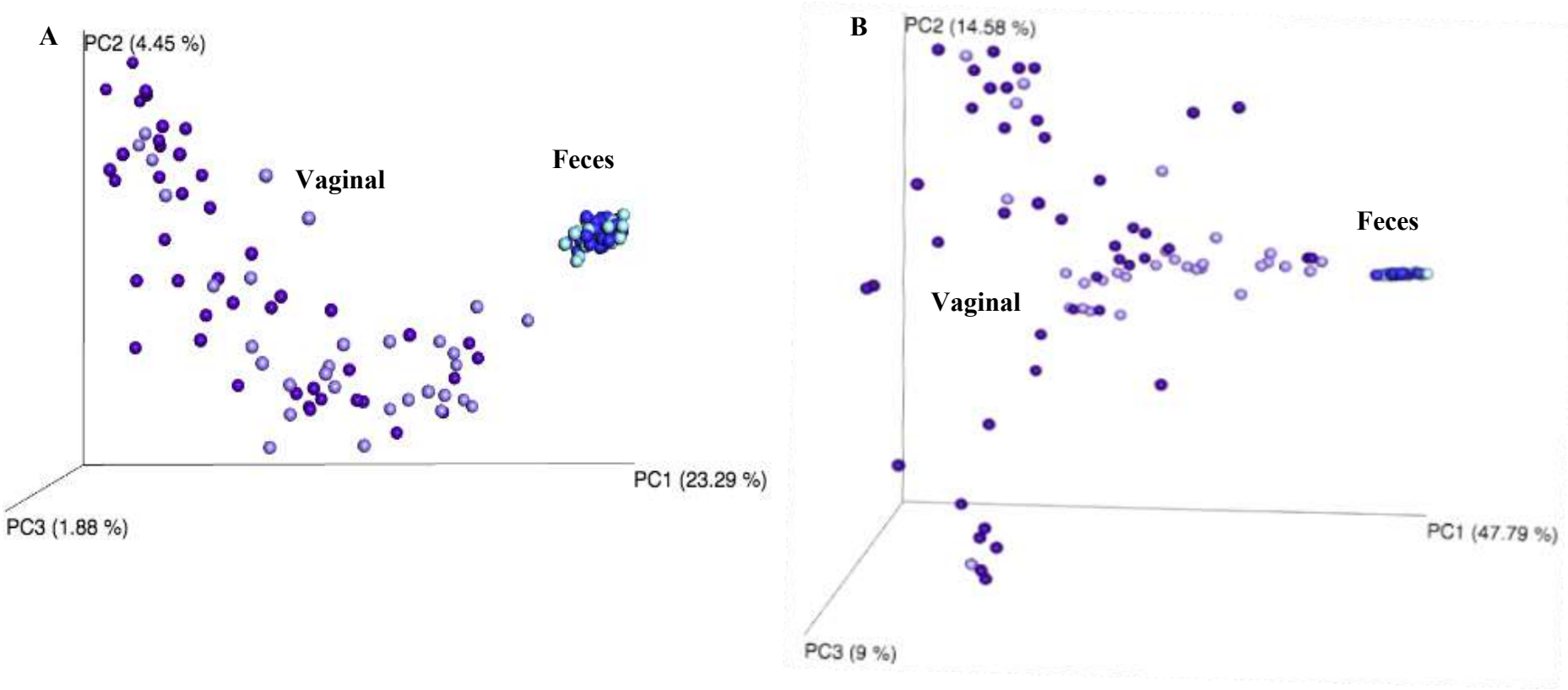
Zand SM (1976). Indexes associated with information theory in water quality. *J Water Pollut Control Fed*: 2026-2031.

Zinicola M, Lima F, Lima S, Machado V, Gomez M, Döpfer D *et al* (2015). Altered microbiomes in bovine digital dermatitis lesions, and the gut as a pathogen reservoir. *PLoS One* 10: e0120504.

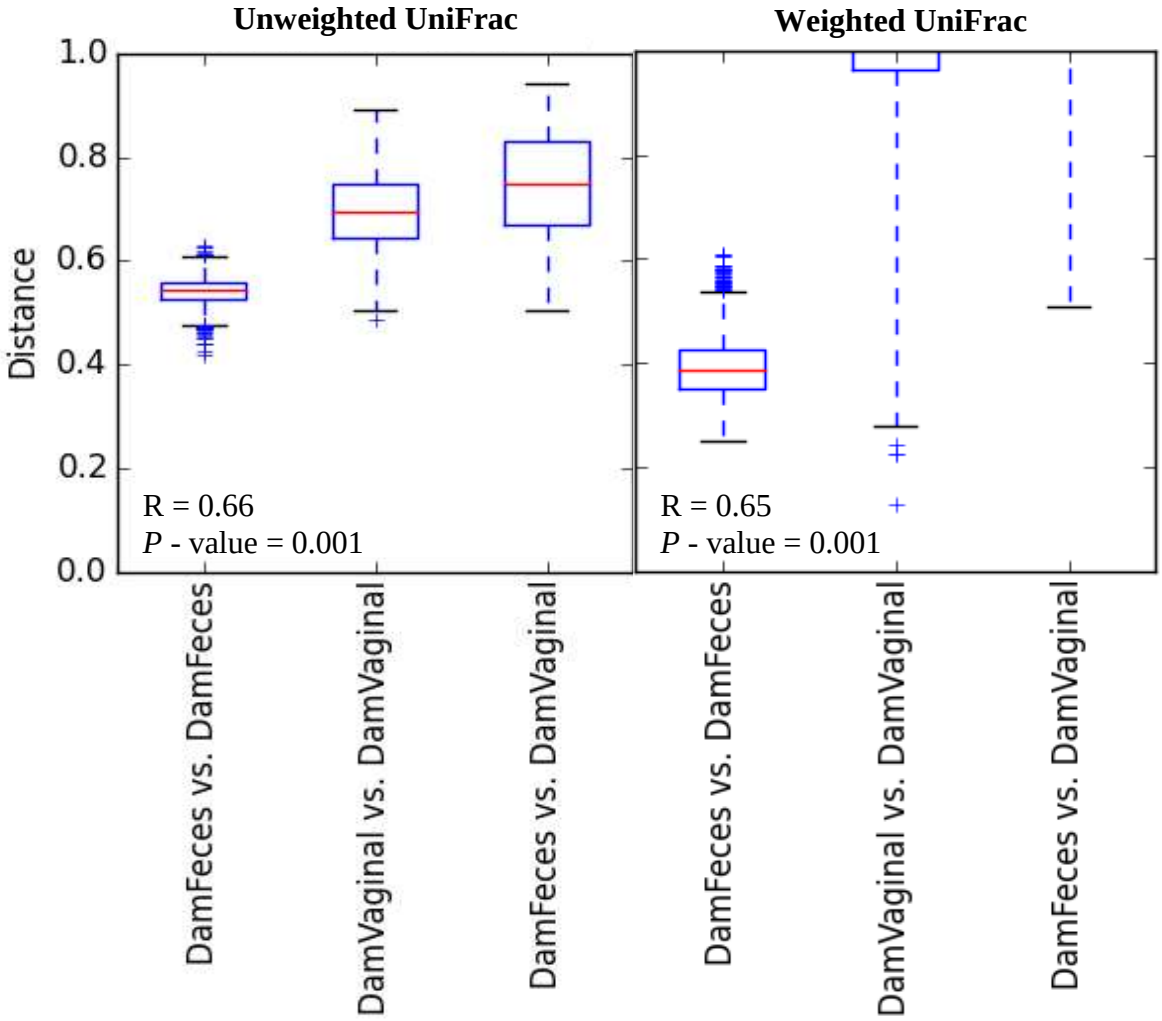
Supplemental 5.1: Chao 1 richness (A) and Shannon diversity (B) indexes according to dam and calf sample types, as well as calf days of life (days 3, 14 and 35). Error bars represent the standard deviation. [a,b,c,d,e] different superscripts among body niches represent a significant difference ($P < 0.05$).



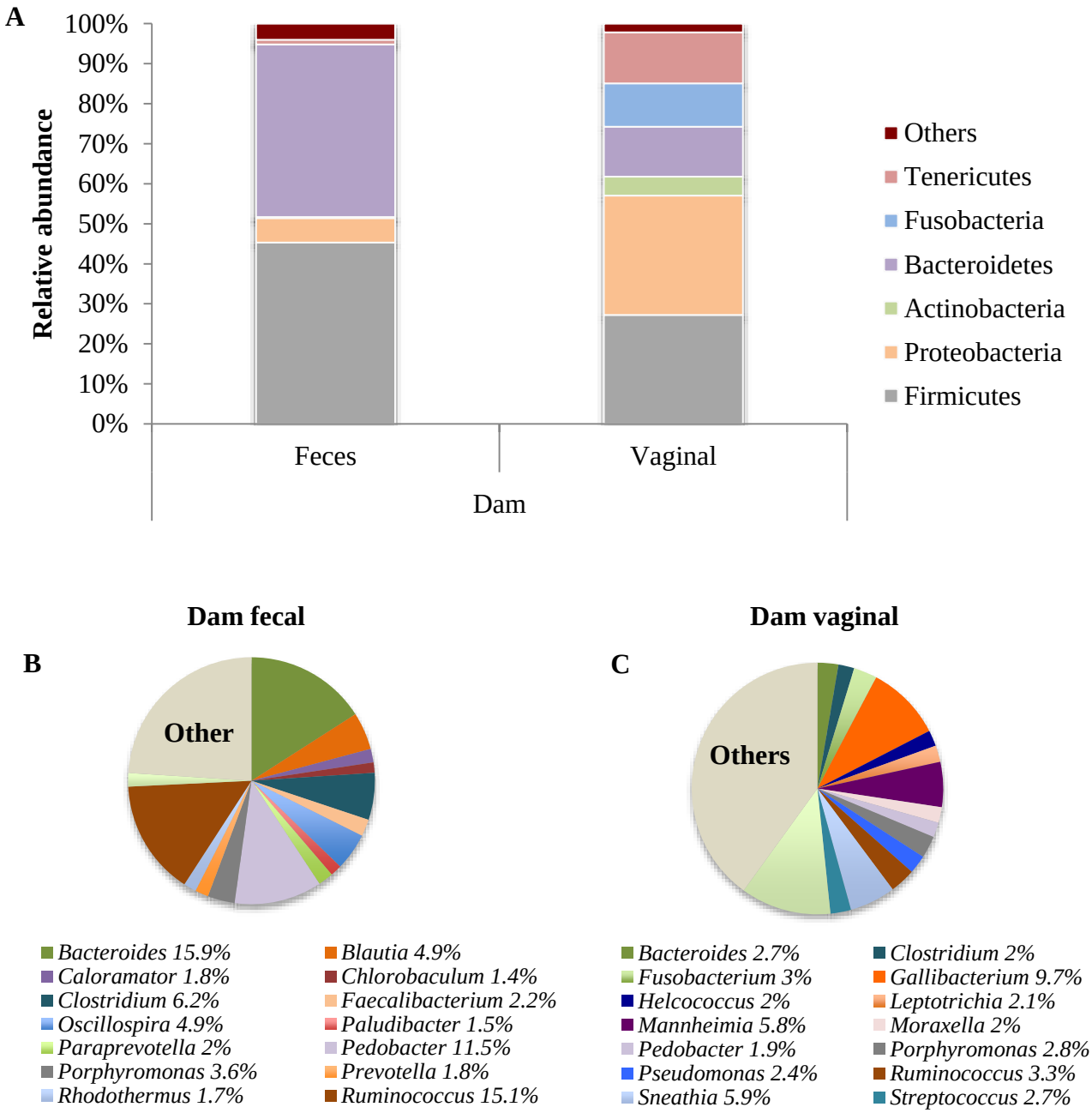
Supplemental 5.2: Principal coordinate analysis (PCoA) of unweighted (A) and weighted Unifrac (B) distances according to cow parity (primiparous and multiparous), and cow body site (gut and vagina). Dark purple dots represent vaginal sites of primiparous cows; light purple, vaginal sites of multiparous cows; dark blue, feces of primiparous cows; and light blue, feces of multiparous cows.



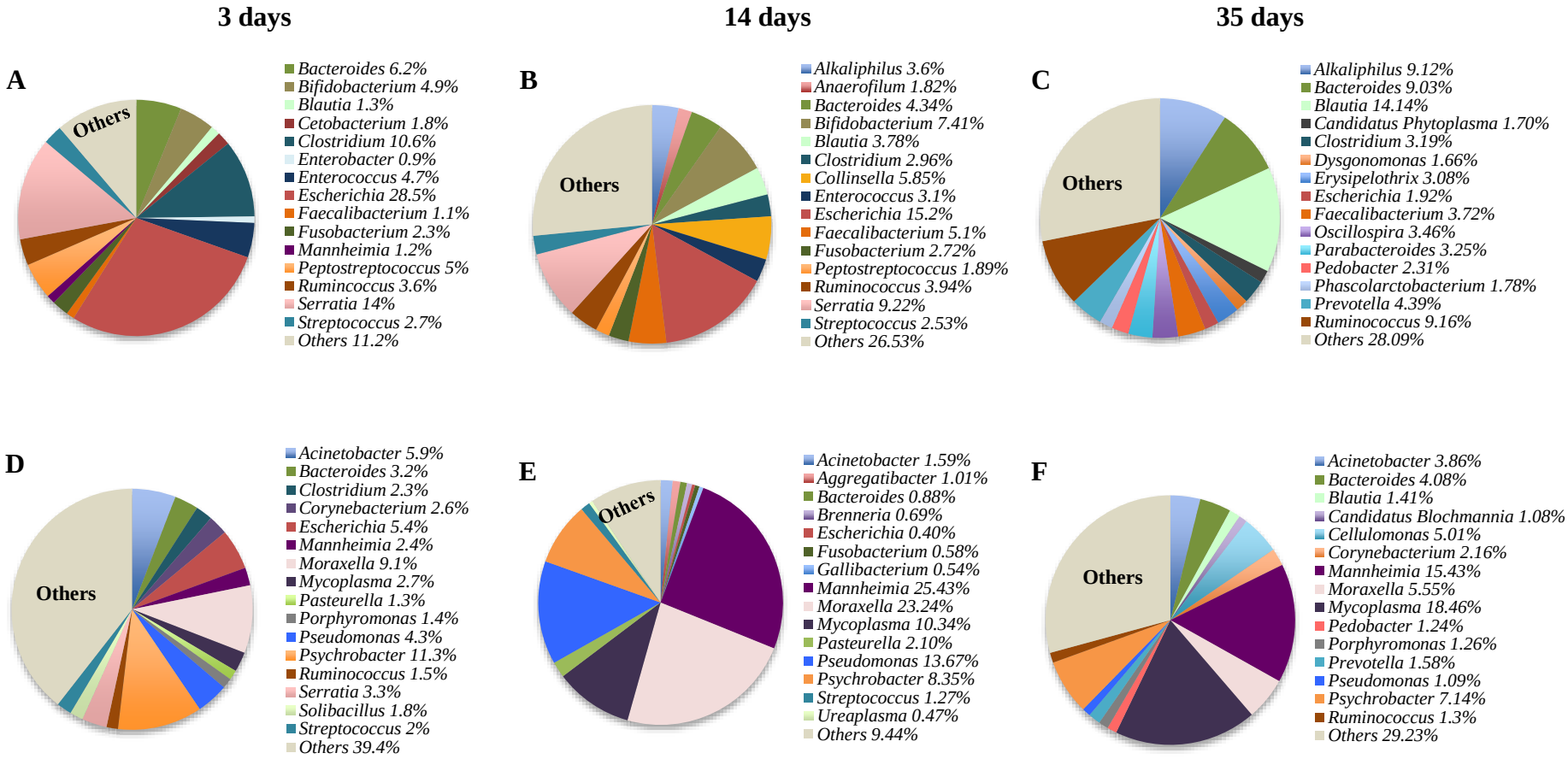
Supplemental 5.3: Unweighted and weighted UniFrac distance differences between dam vaginal and feces microbiotas. An ANOSIM test was performed to compare distances between dam feces and vaginal samples and was based on 999 permutations. The test statistic R from ANOSIM can range from 1 to -1. An R value close to 1 suggests dissimilarity between groups, whereas an R value close to 0 suggests similarity between groups.



Supplemental 5.4: Dam fecal and vaginal microbial composition at the phylum (A) and genus (B and C) levels. Results are shown as the mean relative abundance (MRA) for the dominant phyla (A) and for the 15 most common genera (C and B) detected.



Supplemental 5.5: The 15 most common genera detected in calf feces at days 3 (A), 14 (B) and 35 (C) of life, as well as in the calf URT at days 3 (D), 14 (E) and 35 (F) of life. Results are shown as the mean relative abundance (MRA) for the dominant genera.



CHAPTER SIX

OVERALL CONCLUSIONS AND FINAL REMARKS

The rapid advances in next generation sequencing came accompanied by the plunging costs and higher throughput. Beyond the improvement in technology and cost, they are revolutionizing our understanding of microbial communities in their natural environments and providing a profound knowledge on host-microbiome interactions and their effluence on human and bovine health. The present research explores the use of amplicon sequencing technology to promote and advance the knowledge on the bovine microbiome and its host health.

Isolation of milk metagenomic DNA and its effect on bacterial community profiling and the detection of mastitis causative agents

Since the accuracy of amplicon sequencing in describing the milk bacterial community is a better representative of the original bacterial population and its use as a sensitive tool to diagnose the causative agents of mastitis depends on the efficiency of DNA isolation methods, the research described in Chapter 2 was conducted in an effort to identify the optimum DNA extraction approach that most accurately extracts heterogeneous metagenomic DNA from milk samples of clinical and non-clinical cases of mastitis. The major goal was to improve the efficiency of amplicon sequencing by identifying the DNA extraction protocol and commercial kit that most efficiently isolates the majority of heterogeneous bacterial species of non-mastitic and mastitic milk samples and the method that most accurately diagnoses the causative agent of clinical mastitis.

This research provides primary insights into the effect of DNA isolation methods on DNA concentration, PCR detection efficiency and genomics-based inference of microbiome composition of non-mastitic and mastitic milk samples. Furthermore, it serves as a guide for selecting methods to extract DNA from milk samples from clinical and non-clinical cases of mastitis based on pre-sequencing and post-sequencing outcomes, as well as extraction time and cost. We provide strategies for DNA isolation of milk samples that offers a representative insight into the overall bacterial community composition and which can be performed in a time- and cost-efficient manner.

Considerations

In the research described in Chapter two, 47% of the non-mastitic milk samples failed to be detected by PCR (bands of 390 base pairs size were not detected in the agarose gel), however no significant differences in the percentage of fails were observed between DNA extraction procedures evaluated. Accordingly, isolation of the metagenomic DNA from non-mastitic milk samples still need to be improved and new approaches still need to be generated for isolation of this type of milk sample that typically encounter low bacterial load.

Another point that should be considered is that although the eight DNA extraction procedures (four DNA extraction protocols and two commercial DNA extraction kits) uniformly described the most abundant bacteria of mastitic and non-mastitic milk samples, and efficiently detected the causative agents of mastitis, high disagreement levels, characterized as the number of unique OTUs, were found among DNA extraction kits and protocols. Essentially, the disagreement here observed was due to the ability of

the DNA extraction procedure in detecting rare OTUs (low abundant OTUs at the family level) but not in their ability in detecting the most abundant OTUs. Therefore, precaution when choosing the protocol and kit for DNA extraction of milk from clinical and non-clinical cases of mastitis may need to be taken if the goal of the investigation is to assess and investigate rare bacterial organisms.

Host-microbiome interaction and its effect on host health

Colostrum microbiome and its effect on mammary gland health

The study described in Chapter three, was conducted in order to identify and characterize the bovine colostrum microbiome as well as to elucidate the colostrum bacterial profiles that might be closely associated with mammary gland infection soon after parturition due to the common high incidence of intra-mammary infections in the early fresh period. Identification of new intervention procedures especially in this period may help to improve udder health and milk production.

Based on results of high-throughput sequencing of the 16S rRNA gene, bovine colostrum samples were shown to harbor a diverse microbial community, which was richer (higher number of bacterial species) in primiparous than in multiparous animals and some differences in colostrum taxonomic structure between parities was observed. Additionally, it was found that colostrum harbors some diversity and taxonomic markers of mammary gland health specific to primiparous cows only.

The upper respiratory tract microbiome and its effect on respiratory tract and middle ear health

Given the complex microbial composition of the URT anatomical area, which is comprised of both commensal bacterial organisms and potential pathogens of the lower respiratory tract and middle ear, the study presented in Chapter four was designed to investigate the composition of the bovine URT microbiota and whether/how its changes over time are associated with the development of pneumonia and otitis media during the critical pre-weaning period.

This research showed that while the bacterial community of the URT of newborn calves is not predictive of pneumonia, the total URT bacterial load, as defined by the log of the copy numbers of the 16S rRNA gene quantified by quantitative PCR technique, is a key contributor to the development of pneumonia. A potential new pathogen of the lower respiratory tract and middle ear, the genera *Moraxella*, was also detected in the URT of pre-weaning dairy calves. Furthermore, the results of this study corroborate with the existing knowledge that species of *Mannheimia* and *Mycoplasma* genera are important pathogens of pneumonia and otitis media. As a whole, the present study provides an unprecedented knowledge of the temporal evolution of the bovine URT microbiome in pre-weaning dairy calves as well as its potential contribution to the development of lower respiratory tract and middle ear diseases.

Maternal microbiome effect on offspring

Natural vertical transference of maternal microbes to the neonate, especially via the vaginal canal during birth and its potential effect on offspring health has already been recognized in humans, however the same process has not yet been evaluated in bovines. The research described in Chapter five was conducted in an effort to elucidate the

potential vertical transfer of bacterial organisms from cows to their progenies and the maternal microbial influence on calf health during the pre-weaning period.

The results of Chapter five suggested that the maternal vaginal microbiota is vertically transmitted to the URT of the newborn calf and its influence seems to persist, at least until 35 days of life. A core shared number of OTUs between maternal vaginal microbiota and calf URT microbiota was also observed, and the genera *Mannheimia* and *Moraxella*, which are considered important bacteria for the development of respiratory tract and middle ear diseases in calves, were the most abundant bacteria of this core. Additionally, the genus *Mannheimia* was relatively more abundant in the vaginal microbiota of dams whose progeny did develop URT disease. Together, these findings indicate that the health status of the calf's initial respiratory tract and middle ear might be affected by mother-to-offspring transmission of bacterial organisms. This study also revealed that the gut (as represented by feces) microbiota of the newborn calf differed from the maternal gut; however, the calf gut microbiota change over time, and by day 35 of life it resembled the microbiota observed in the mature bovine gut.

This research provides pioneering knowledge on the existence of the vertical transference of maternal microbial organisms to the *Bos taurus* neonate as well as its potential contribution to the health of the offspring respiratory tract and middle ear.

Considerations

As documented in Chapter five, experiments in humans unveiled the potential for transfer of bacterial communities from the mother to her progeny via other maternal-routes excluding the vaginal canal. For instance, maternal antibodies that are passed in

milk to the progeny were described as a potential influencer of the progeny gut microbial composition. Furthermore, studies have shown that the placenta and meconium are not sterile habitats. Therefore, despite the evidence of the maternal microbial impact on offspring initial microbial colonization and its effect on offspring health described in Chapter five, further studies are still needed to confirm the maternal protective effect on offspring health and evaluate to what extent the microbiota composition detected in the offspring URT was due to transfer of the maternal vaginal microbiota at birth.

Final remarks

Fast and accurate identification and characterization of disease causative bacterial agents are crucial for successful treatment, animal recovery and disease control. In the past two decades molecular diagnostic methods have advanced substantially and rapidly. The research described in this dissertation was novel and demonstrated the efficiency of amplicon sequencing as a diagnostic tool. Additionally, it supports the growing recognition of the importance of the host-microbiome interaction and its contribution to bovine-host health. Together, the results of this dissertation provided unprecedented knowledge on the bovine microbiome and its potential effect on animal health.