IDENTIFYING BOVINE RESPIRATORY DISEASE (BRD) THROUGH THE NASAL MICROBIOME

by

Ruth Eunice Centeno Martinez

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Animal Sciences

West Lafayette, Indiana

May 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Timothy Johnson, Chair
Department of Animal Science

Dr. Jacquelyn P Boerman
Department of Animal Science

Dr. Jennifer Koziol
Department of Veterinary Clinical Science

Dr. Esteban Fernandez-Juricic
Department of Biological Sciences

Approved by:
Dr. John Blanton

To the Almighty God, family and friends for their unconditional love, support and guidance

ACKNOWLEDGMENTS

The author is very grateful to God Almighty for His guidance and wisdom He bestowed upon me, for the strength, peace of mine, and good health that allowed me to finish this research.

I would like to express my gratitude towards my family and friends back in El Salvador and Honduras, for the encouragement which helped me to achieve this goal. Papi, Mami, and Samu, my family, to which I deeply dedicate this work, thanks for your love and words of encouragement that gave me strength throughout this process.

I would like to express my gratitude to Dr. Timothy Johnson, my research advisor. Thanks for teaching me, guiding me, and introduce me to the wonderful world of Microbiome, it has been an honor to be a member of your lab, and thank you for always be willing to help me no matter how busy you were. I would also thank my lab mates to which I have the honor to call my friends, Carmen, Paul, and the wonderful undergrads Natalie, Johnie "Junior", Tessa, Jenn, and Tiffany for their incredible work. In addition, I would like to thank my committee members Dr. Jennifer Koziol, Dr. Jacquelyn Boerman, and Dr. Esteban Fernandez-Juricic whose guidance allowed me to fulfill this incredible goal in my life; thanks for teaching me, it was an honor to learn from you. Also, a thank you to my friends from the Zamorano Alumni Association at Purdue and the Purdue Chi Alpha Christian Fellowship for their support throughout this process.

Lastly, a thank you to Dr. Mohit Verma and his students for leading the Bovine Respiratory Disease research project. We would like to thank the US Department of Agriculture – National Institute for Food and Agriculture (Grant # 2020-68014-31302) and Purdue AgSEED who provided funding that made this work possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

BRD:	Bovine Respiratory Disease
ASV:	Amplicon Sequence Variants
PCR:	Polymerase Chain Reaction
qPCR:	Real-time Quantitative Polymerase Chain Reaction
DNA:	Deoxyribonucleic Acid
rRNA:	Ribosomal Ribonucleic Acid

ABSTRACT

Bovine respiratory disease (BRD) is an ongoing health and economic issue in the dairy and beef cattle industry. Also, there are multiple risk factors that make an animal susceptible to BRD and it's diagnosis and treatment is a challenge for producers. Four bacterial species, *Mannheimia haemolytica, Pasteurella multocida, Histophilus somni,* and *Mycoplasma bovis* have been associated with BRD mortalities. Hence, this study aims to characterize the cattle nasal microbiome as a potential additional diagnostic method to identify animals suspected to have a lung infection. Quantitative PCR and 16S rRNA gene sequencing were used to determine the bacterial load of these four bacterial pathogens in the nasal microbiome of apparently healthy (N=75) and (N=58) affected by BRD Holstein steers. We then sought to identify a value or equation that could be used to discriminate between BRD and healthy animals using a Linear Discriminant Model (LDA). Additionally, co-occurrence between commensal bacterial and BRD-pathogens were also identified. Cattle diagnosed with BRD presented lower richness, evenness and phylogenetic diversity than healthy pen-mates. Bacterial species and genera *Truperella pyrogenes* and *Bibersteina* were increased in the BRD group, and the species *Mycoplasma bovirhinis* and *Clostridium sensu stricto* increased in the healthy group. Prevalence of *H. somni* (98%) and *P. multocida* (97%) were the highest regardless of disease diagnosis in all the samples. Prevalence of *M. haemolytica* (81 vs. 61%) and *M. bovis* (74 vs. 50.7%) were higher in the BRD group. The bacterial density of *M. haemolytica* and *M. bovis* was also higher in the BRD group, whereas *Histophilus somni* was lower in the BRD group. Five different models were tested using LDA, and one model produced a sensitivity and specificity of 60% and 81% agreement with diagnosis based on animal symptoms. Co-occurrence analysis demonstrated that the nasal microbiome members are more likely to interact with each other than associations between BRD-pathogens and nasal microbiome members. This study offers insight into the BRD-pathogens prevalence and difference in nasal microbiome between healthy and BRD animals and provides a potential platform for future studies and potential pen-side diagnostic testing.

CHAPTER 1. INTRODUCTION

1.1 Background and Significance

Bovine respiratory disease (BRD) is an ongoing health issue in the dairy and beef cattle industries. BRD treatment represents a high expense to producers, and it's been estimated that the annual BRD expenses in the US to be $800-900 M (Buckham et al., 2008; Chirase & Greene, 2000). The development of this disease is associated with multiple risk factors (i.e., predisposing, environmental, and epidemiological factors) that make the animal susceptible to BRD (Snowder et al., 2006). In the cattle industry, the most common way to diagnose an animal with BRD relies on visual observations of animal behavior and measuring rectal temperature (Griffin et al., 2010). Nevertheless, sometimes a sick animal will not present common clinical signs because of the predatory/prey behavior in which they mask early symptoms; therefore, it cannot be identified or treated (Duff & Galyean, 2007; Timsit et al., 2011). Because of this, different methods that include white blood cell counts, thoracic ultrasonography, and others, had been studied to identify and classify a sick animal from a healthy animal (Wolfger et al., 2015; Schaefer et al., 2007, 2012; Al-Ani et al., 2015). Unfortunately, the results in BRD diagnosis are inconsistent and more research is needed to quickly and accurately identify animals in need of antibiotic therapy.

As previously discussed, BRD can be caused by multiple factors such as predisposing, environmental, and epidemiological factors. Among the epidemiological factors, studies have identified the bacterial pathogens *Mannheimia haemolytica, Pasteurella multocida, Histophilus somni*, and *Mycoplasma bovis* are related to BRD (Klima et al., 2014; Mosier, 2014). These bacteria are considered commensal microbes in the respiratory tract, but under stressful conditions (e.g., weaning, transport, or stress), they can become pathogenic and cause BRD (Mosier, 2014). Because of this, different molecular methods, including 16S rRNA gene sequencing and qPCR, had been developed to characterize and quantify the bacteria present in the cattle's respiratory tract between healthy and sick cattle (Pardon & Buczinski, 2020). Even though the use of molecular techniques can be utilized for the detection and measurement of the specific bacterial load in a sample, the issue arises when the BRD-associated bacteria are present in healthy and BRD animals, making challenging to differentiate pathogenic from commensal bacteria (Pansri et al., 2020).

Also, it is known that within the microbial community, bacteria are in constant competition to survive in the environment (Corbeil et al., 1985). Different studies reported that bacteria like *Mycoplasma dispar, Lactococcus lactis,* and *Lactobacillus casei* were significantly higher in healthy animals than BRD-animals. In contrast, sick animals present a higher abundance of *Mycoplasma bovis, Mannheimia haemolytica,* and *Pasteurella multocida* (Timsit et al., 2018). Therefore, these differences in the microbiome composition between healthy and sick animals could reveal bacteria-bacteria relationships occurring in the cattle respiratory tract between commensal and pathogenic bacteria during disease and why some animals do or not develop BRD.

1.2 Objectives

Objectives for this study are:

Objective 1: Perform a 16S rRNA gene amplicon sequencing using the DNA extracted from nasal swabs to characterize the microbial community present in the nasal cavity between healthy and BRD pen-mates.

Objective 2: Perform a qPCR reaction using the extracted DNA to target and quantify the relative abundance of *Mannheimia hemolytica*, *Pasteurella multocida, Histophillus somni, Mycoplasma bovis* and 16S rRNA gene in the nasal cavity of healthy and BRD pen-mates to identify a possible equation or model that could be used to discriminate BRD from healthy animals.

Objective 3: Perform a co-occurrence to identify species pairs that are common among BRD and healthy groups and any co-occurrence between BRD-pathogens and nasal microbiome.

1.3 Organization of the Thesis

This thesis is organized following the Purdue Graduate School traditional style format. Chapter 2 provides a literature view about the BRD background, and molecular techniques used to identify the bacteria in the respiratory tract from healthy and those diagnosed with BRD. Chapter 3 provides the material and methods used in conducting the experiments. Chapter 4 provides the experimental results. Lastly, chapter 5 contain the discussion of the results and summarizes the important

findings in the study as well as future research ideas that can be perform with the knowledge obtained by this research.

CHAPTER 2. LITERATURE REVIEW

2.1 Bovine Respiratory Disease

Bovine respiratory disease (BRD) is an ongoing and vital problem in dairy and beef cattle. It is defined as an infectious pneumonia, which can cause pulmonary lesions (Muggli-Cockett et al., 1992). BRD affects the health of cattle of all ages, including feedlot, dairy calves, nursing beef cattle, post-weaned cattle by compromising its immune system and causing morbidity and mortality (Babcock et al., 2010; Holman et al., 2015a). BRD can account for different ranges of illnesses such as 1) peracute, which account for animals with acute fibrinous pneumonia, 2) acute, which refers to active pneumonia lesion, 3) subacute, which refers to chronic and active pneumonia lesion 4) bronchiolar, chronic and active pneumonia lesion with "bronchiolar" pattern, and 5) chronic pneumonia, animals who received at least 3 BRD treatments (Booker et al., 2008).

2.2 BRD economic impact in the dairy and beef industry

Bovine respiratory disease represents a critical economic problem to farms (Buckham et al., 2008). The economic impact of a BRD treatment has been estimated to be $800-$900 M related to animal death, reduction of feed efficiency, and treatment costs (Chirase & Greene, 2000). It has been reported that BRD is responsible for approximately 75% of the morbidity and 57% of mortality in the feedlots (Vogel et al., 2015). In dairy cattle, BRD cause 24% of death in pre-weaned heifer and 58.9% death in weaned heifers (USDA, 2014). Also, 89% of sick cattle with BRD are treated with injectable antibiotics, but in 33% of the cases, the treatment fails, resulting in additional treatment is required or animal death (Avra et al., 2017). A study reported that producers lose $40.46/calf for one BRD treatment, $58.35/calf for two treatments, $291.93/calf for three or more treatments from the net value (carcass value- total cost in feedlot) (Fulton et al., 2002).

2.3 BRD development

2.3.1 Factors that cause BRD

Different studies had identified that BRD could be developed by the action of multiple predisposing, environmental, and epidemiological factors that make it harder to diagnose and treat

a sick animal. The predisposing factors include animal age, handling stress, and transport from one farm to another. The environmental risk factors include stocking density, ambient temperature, humidity, and ventilation (Edward, 1996; Snowder et al., 2006). The epidemiological factors include bacterial and viral agents, the interaction between viruses, mode of transmission, as well as the infectious, latent, and carrier periods (Bowland & Shewen, 2000; Edward, 1996; Snowder et al., 2006). Common BRD-associated bacterial species are *Mannheimia haemolytica, Pasteurella multocida, Histophilus somni* and *Mycoplasma bovis* (Klima et al., 2014; Mosier, 2014).

2.3.2 Disease development

The animal's respiratory system possesses different line of defenses (e.g. nasopharynx mucosal layer, mucosal epithelium and network signaling and communication). These line of defenses produce antimicrobial peptides, glycoproteins, IgA proteins, luminal and mucosal surface macrophages and dendritic cells that regulates microbial homeostasis between commensals and pathogenic bacteria, clearance of potential pathogens and recruit immune cells to protect the animal's health (Ackermann et al., 2010; Uehara et al., 2007; Zeineldin et al., 2019). In addition to the animal susceptibility, it has been reported that viruses present in the respiratory tract predispose animals to bacterial infections. The viral infection can damage the respiratory and lung parenchyma, which facilitates the translocation of bacterial pathogens and the viral infection causes a delay in the animal immune response to the bacterial infection (McMullen et al., 2019; Taylor et al., 2010; Timsit et al., 2016). A healthy animal presents a stable community, in which the microbes are compartmentalized and neutralized within the lumen by the antimicrobial peptides (Zeineldin et al., 2019). Nevertheless, when the animals suffer from stress caused by multiple predisposing and environmental factors, these defenses fail making them susceptible to BRD (Caswell, 2014; Timsit et al., 2016). Once a distribution in the stable state occurs in the respiratory tract, the pre-disease state begins in which there is a loss of mucosal barrier function as a result of bacterial dysbiosis and pathogen colonization in the epithelium; Nonetheless, at this stage, with the action of proinflammatory cytokines, chemokines and activation of local immune cells, the health of the animals can be restore. However, if the disease progress, the mucosal barrier is lost which results in microbial translocation across the mucosal epithelium and failure in immune regulation (Zeineldin et al., 2019). Once the bacterial pathogens move to the lungs, they

adhere and colonize the epithelial surface, provoking pulmonary inflammation, and gross pathology (Caswell, 2014).

2.3.3 Animal Susceptibility

In feedlots, recently weaned, transported, non-vaccinated, lightweight, commingled, and auction-market derived cattle are considered high-risk cattle to suffer BRD (Holman et al., 2019; Timsit et al., 2017). Snowder (2006) analyzed BRD incidence in calves from 9 different cattle beef breeds over a 15-yr period (1987-2001) and identified that the BRD incidence was significantly different between heifers and steers. In the study, steers reported 20% of BRD incidence compared heifers in which only 14% had BRD. Also, Snowder (2006) reported that castration before feedlot entry could cause male calves to be more susceptible to developing BRD.

Another critical factor that can increase the animal susceptibility to BRD is the failure of passive transfer (FPT) (Gorden & Plummer, 2010). FPT refers to the failure in the absorption of immunoglobulin (Ig) from the consumption of colostrum by the calf after birth and can increase mortality until week 10 of the animal age (Hogan et al., 2015; Tyler et al., 1999). Colostrum consumption is crucial to the calves because it provides passive immunity, which reduces the animal's morbidity and mortality (Roy, 1980). Thus, providing colostrum to newborn calves might increase the animal's immunity and decrease the development of diseases like BRD at a young age.

In addition, it has been reported that cattle at parturition and lactation state might play a role in the offspring from older animals to suffer from BRD (Gorden & Plummer, 2010). This is due to the negative energy balance of the cow, decreased dry matter intake, nutrient absorption, as well as health conditions like ketosis or hypocalcemia during lactation and parturition (Goff & Horst, 1997; Kimura et al., 2006; NRC, 2001).

2.4 Detection of animals with BRD

Most farmers rely on animal behavior and appearance observations to diagnose cattle with respiratory disease. The most commonly targeted clinical signs are depression, appetite loss, respiratory character change, and increased rectal temperature, commonly referred to as the DART

method (Griffin et al., 2010). Nevertheless, the observation of clinical signs is often not accurate to diagnose an animal with BRD, evidenced by the method's low sensitivity (62%) in correctly identifying individuals with the disease and low specificity (63%) in identifying animals without the disease (Loong, 2003; White & Renter, 2009; Wolfger et al., 2015). For example, Timsit et al. (2011) indicated that 74% of 449 cattle fever episodes were not visually detected by the feedlot personal. Due to the inconsistency in BRD signal detection, cattle with BRD are often detected late or not detected at all (Timsit et al., 2011). Because of this, different methods had been studied to identify truly sick animals with BRD from a healthy animal. A group measured the acute phase proteins in the animal's blood and identified that animals with BRD presented a higher concentration of haptoglobin (HP) and lipopolysaccharide-binding protein (LBP) in the blood than healthy animals (Idoate et al., 2015). In the same study, a concentration of ≥0.81 mg/ml of HT (method sensitivity and specificity: 92.86 and 85.71%) and ≥ 0.33 μg/ml of LBP (method sensitivity and specificity: 92.86 and 92.86%) in the blood were proposed to be used as an indicator of BRD.

Also, white blood cell counts (WBC) and neutrophil/lymphocyte ratio have been used to detect BRD (Schaefer et al., 2007, 2012). In these studies, the WBC sensitivity and specificity were between 25% and 77.8% and between 77.4% and 94%. For the neutrophil/lymphocyte ratio, the sensitivity and specificity were 38%-66.7% and 67.9%-86.5. Nevertheless, due to the wide sensitivity and specificity results, WBC and neutrophil/lymphocyte ratio have limited value in differentiating truly BRD animals from healthy animals. Another method used to detect BRD is thoracic ultrasonography. With thoracic ultrasonography, it is possible to determine the lung lesions and consolidation of the animal that refers to the pathological process when the alveoli are full with fluids, pus, blood, cells, protein, and not air (Al-Ani et al., 2015). Ultrasonography detected lung lesions at least one time during the study in 28% of the BRD animals, in 16% of the control groups (Abutarbush et al., 2012). Also, in a different study, lung lesions identified with Ultrasonography presented a sensitivity and specificity 94 and 100% in cases associated with BRD (Ollivett et al., 2015). A new method being studied is metabolomics that measures small molecules in cells, tissue, and biofluids and can be used as biomarkers to indicate disease (Goldansaz et al., 2017; Moore et al., 2007). The combination of blood metabolome profile with BRD visual diagnosis and visual-clinical diagnosis had sensitivity between 82-88%, a specificity between 74-

87%, and accuracy between 81-85% in BRD detection (Blakebrough-Hall et al., 2020). In the same study, the metabolites identified that help classify the animals as BRD or non-BRD were tyrosine, citrate, and hydroxybutyrate, as well as some unknown metabolites. Although multiple methods had been tested to diagnose BRD, further evaluation is needed to confirm the sensitivity, specificity, and accuracy of each method in BRD diagnosis.

2.5 BRD management practices

2.5.1 Vaccination

Dairy and beef producers vaccinate their cattle to enhance animal immunity to pathogens by increasing the antibody concentration in the body (Wilson et al., 2017). In the U.S, two thirds of feedlots vaccinate for common BRD-bacterial agents *Histophilus somni, Mannheimia haemolytica,* and *Pasteurella multocida* as means of BRD prevention (NAHMS, 2013). Producers may use modified-live viral vaccines (MLV), killed virus (KV) vaccines or a combination of BRD-associated bacteria bacterin/toxoids that are commercially available against BRD-associated viruses and bacteria (Chamorro & Palomares, 2020; Edwards, 2010). The benefits of MLV vaccines include a long-lasting immune response, fewer doses, less reliance on adjuvants, vaccine antigens that resemble the pathogenic organism and stimulated the effector component of the cell-mediated immunity. The benefits of KV vaccines include longer shelf life, reduced virulence reversion, storage stability, and less likely to be contaminated with another organism (Edwards, 2010). In a study, cattle at feedlot arrival were divided into two different vaccination programs. Group one were vaccinated with MLV containing two BRD-associated viruses: infectious bovine rhinotracheitis virus (IBRV), types I and II bovine viral diarrhea virus (BVDV), and *Mannheimia haemolytica* with *Pasteurella multocida* bacterin-toxoid whereas group 2 received a vaccine containing only IBRV, type I BVBV, bovine respiratory syncytial virus, parainfluenza-3 virus, and *Mannheimia haemolytica* bacterin-toxoid, a less complex vaccine. Animals in group one presented a significant decrease in BRD treatment, overall chronicity defined as animals with chronic disease divided by the total number of animals present in the study, wastage defined as animals considered chronic but did not die divided by the number of animals present in the study, mortality, and BRD-specific mortality than the animals that received the second vaccination program (Wildman et al., 2008). Also, studies demonstrated that vaccination of beef calves during

weaning with MLV alone or with *Mannheimia haemolytica/Pasteurella multocida* bacterins could reduce BRD morbidity and mortality after weaning. However, the vaccination effect in young beef and dairy calves with MLV or with *Mannheimia haemolytica/Pasteurella multocida* bacterins is uncertain, or there is conflicting evidence of its efficacy (Chamorro & Palomares, 2020). Another study summarized that the vaccination efficacy against *Mannheimia haemolytica, Pasteurella multocida,* and *Histophilus somni* is inconsistent (Griffin et al., 2010). The inconsistency of the vaccination effect has been attributed to 1) some studies with natural BRD occurrences tend to mix the vaccinated and non-vaccinated calves in the same pen, which underestimate the effect of vaccination, 2) other studies report the crude mortality and morbidity of the pen, while others report only BRD-mortality and morbidity (Larson & Step, 2012). Therefore, it is complicated to determine if vaccination against BRD-associated pathogens may reduce BRD-specific morbidity and mortality.

2.5.2 Antibiotic treatment

In the feedlot industry, the use of antibiotic therapy represents a vital management activity to control and treat BRD (Holman et al., 2019; Watts & Sweeney, 2010). Metaphylaxis, which corresponds to the timely mass medication of a group of animals to eliminate or decrease the outbreak of a disease, has been used in the feedlots to control bacterial pathogens related to BRD outbreaks (Edwards, 2010; Zeineldin et al., 2020a). The injectable antimicrobials ceftiofur crystalline free acid, florfenicol, oxytetracycline, tilmicosin, tulathromycin, as well as feed grade antimicrobials such as chlortetracycline (CTC) and chlortetracycline plus sulfamethazine are approved to be used for metaphylaxis in the feedlots (Edwards, 2010; Holman et al., 2019). Additionally, diagnosis of an animal with BRD is sometimes challenging to producers making the efficacy of antibiotic treatment to be inconsistent due to infection misdiagnosis and ineffective use of antibiotics as prophylaxis, metaphylaxis and growth promoters (Ives & Richeson, 2015). One concern that has arisen with the antibiotic treatment is the antibiotic resistance in BRD pathogens (Call et al., 2008; Griffin et al., 2010; Rice et al., 2008). BRD associated pathogens had been identified to be resistant to antibiotic like tetracyclines, fluoroquinolones, beta-lactams, macrolides, sulfonamides, lincosamides, phenicols, and aminoglycosides (Dedonder & Apley, 2015).

Antibiotic administration is associated with decreased bacterial richness and disruption in the respiratory microbial community structure because of their bactericidal or bacteriostatic effect on pathogenic or commensal microbes or both (Zeineldin et al., 2020a; Zhang et al., 2019). Holman et al. (2019) demonstrated that a single injection of oxytetracycline and tulathromycin caused a perturbation in the bovine nasopharyngeal (NP) and fecal microbiota at 2 and 5 days after giving them the single injection. Also, animals that received the single injection of oxytetracycline had an increase in genes conferring resistance to tetracycline, sulfonamide, and erythromycin: *erm*(X), *sul*2, *tet*(H), *tet*(M) and *tet*(W) in NP samples and *tet*(M) and *tet*(W) in the fecal samples on day 12 after receiving the single antibiotic injection; whereas animals that received the tulathromycin injection had an increase of *erm*(X), *sul*2 and *tet*(M) 34 days after the injection in the NP samples (Holman et al., 2019).

2.6 Molecular methods used to identify BRD-associated pathogens as a means of diseases diagnostic

As previously described, BRD can be caused by epidemiological agents such as the species *Mannheimia haemolytica, Pasteurella multocida, Histophilus somni,* and *Mycoplasma bovis* (Dabo et al., 2008; Griffin et al., 2010; Klima et al., 2014) Because of this, different studies had focused on the use of molecular methods such as microbial culture, single polymerase chain reaction (PCR), real-time quantitative PCR (qPCR), and bacterial 16S rRNA gene PCR amplicon sequencing to identify the BRD-associated bacteria in bovine nasal, nasopharyngeal and lung tissue samples (Pardon & Buczinski, 2020).

2.6.1 Microbial culture

Microbial culture can be used to isolate and identify the bacteria from nasal, nasopharyngeal swabs, transtracheal wash, or lungs necroscopy (Fulton & Confer, 2012). In a study performed by Seker et al. (2009), nasal swabs from 100 healthy and sick animals with nasal discharge, cough and dyspnoea and 220 bacteria were isolated using a Columbia blood agar culture plate. In this study, the authors could isolate *Staphylococcus epidermis* (32.9%) and *S. aureus* (24.3%) on 70 healthy animals, whereas, in the bacterial cultures from 30 unhealthy animals, they identified *Pseudomonas aeruginosa* (40.0%), *P. multocida* (40.0%), and *M. haemolytica* (100.00%). *Mycoplasma bovis* was not identified in this study, which is considered a

BRD-associated bacteria. Parker et al. (2018) mentioned that *Mycoplasma* species have a simple structure and cannot synthesize fatty acids, and because of this, it requires a specific and highly enriched media to grow. When using microbial cultures, fastidious growers like *Histophilus somni* can be quickly overgrown, causing false negatives (Pardon & Buczinski, 2020). These results indicate that depending on the type of media selected to culture the microbes, false negative over positive results are common. One disadvantage of using culture-based methods is that even though it is possible to identify BRD-associated pathogens, a relatively small fraction of bacteria can be cultivated (Rappé & Giovannoni, 2003). Additionally, the sensitivity and specificity of the microbial culture for some BRD-associated pathogens had not been determined; thus, it's not possible to determine the presence of other BRD-associated pathogens (Pardon & Buczinski, 2020).

2.6.2 Microbial culture and Polymerase Chain Reaction (PCR)

A combination of culture-based methods and 16S rRNA gene amplification has been performed to enhance the detection of BRD-associated pathogens from the cattle respiratory tract. Polymerase Chain Reaction is considered a common method used to identify BRD-associated bacteria due to the capacity to detect culturable and unculturable bacteria and viruses (Dickson et al., 2014; Pardon & Buczinski, 2020). For this method, specific primers are needed to identify the pathogens of interest. In most studies, the 16S rRNA gene is used to identify the BRD-associated bacteria or characterize the microbial community. Holman et al., (2015b), collected nasopharyngeal samples of apparently healthy cattle at entry and during feedlot placement, plated the samples in Brain Heart Infusion agar (BHI), de Man, Rogosa and Sharpa agar (MRS), and 5% sheep blood agar followed by the extraction of DNA from the isolates and PCR amplification of near the full length of the 16S rRNA gene. By combining the culture-based method and PCR amplification, the authors could identify from the nasopharyngeal samples the BRD-associated genera *Pasteurella* and *Mannheimia* during feedlot entry and after 60 days, by culturing samples on the BHI and blood media while on the MRS media, the genera *Lactobacillus, Enterococcus, Rummeliibacillus*, and *Pediococcus* were identified in the animal nasopharyngeal samples. Again, no members of the genera *Mycoplasma* and *Histophilus* were identified using the culture-based methods.

In another study performed by Bell et al., (2014), DNA was extracted from 150 samples of cattle pneumonic lung tissue. PCR assays were performed to target *M. haemolytica, H. somni,* and *P. multocida*, then, the results were compared to culture-based methods to detect BRD-associated in the samples. In the study, the authors detected *M. haemolytica* in 51 (34%) cases by PCR and in 33 (22%) cases by culture-based methods, *H. somni* was detected in 35 (23.3%) samples using PCR than the 6 cases (4%) detected by culture, *M. bovis* was detected in 53 (35.3%) cases using PCR, and in 29 cases (19.3%) by culture; lastly, *P. multocida* was detected in 42 (28%) cases using PCR and only in 31 (20.7%) cases by culture. These results indicated that culture-based methods and PCR amplification can help detect the BRD-associated pathogens; however, the results depend on the type of sample collected and the type of growth media selected.

2.6.3 Conventional PCR

PCR gene amplification is a common method selected for BRD-associated pathogen detection. The principle behind PCR focused on the amplification of a known generic region. The most common gene used to characterize a microbial community is the bacterial 16S rRNA gene due to its presence in almost all bacteria (Janda & Abbott, 2007). Nevertheless, PCR can also be used to amplifying the genes that are specific to a bacteria group. One advantage of using PCR is capturing more information about the microbial community composition than culture-based methods. PCR amplification using specific primers increases the sensitivity and specificity for detecting the bacteria of interest. (Pardon & Buczinski, 2020). Nevertheless, once the primers are designed, it is a time consuming and difficult process to validate the primers that could be used to detect the desired bacteria (Thanthrige-Don et al., 2018).

One disadvantage of conventional PCR is the incapacity to measure the bacterial density in the samples. The products from the PCR are compared to a known positive and negative controls using gel electrophoresis. This step serves as a quality point because it indicates if the desired fragment was amplified and if the bacteria is present or not, but it does not specify the bacterial load for each of the targets. A problem arises when BRD-associated bacteria are present in healthy and sick cattle, but there is no accurate way to measure the difference in bacterial density between them (Pansri et al., 2020).

2.6.4 Real-time quantitative PCR (qPCR)

Real-time quantitative PCR (qPCR) can detect and quantify the bacterial species' density in the samples. What makes qPCR able to measure bacterial density is the inclusion of a fluorescent probe at the 5' end and a quenching dye at the 3'end (Jerome, 2010). The principle focused on that during amplification, specimens with a higher density in the sample will amplify sooner than specimens will a low bacterial load. The amplification is detected by the fluorescent signal (or earlier threshold cycle "Ct"). Using the threshold cycle value, it is possible to quantify the number of amplicon copies per mL (Fulton & Confer, 2012). Thomas et al. (2019) performed a Real Time TaqMan PCR (qPCR) assay to target the presence and measure the BRD-associated bacteria's carriage from 299 cattle's nasal swabs over time. The group designed specific primers to target the genes *sodA* and the 16S rRNA gene. *Pasteurella multocida* was detected in 227/299 swabs (75.9%), *Histophillu somni* was detected in 80/299 swabs (26.8%), and *Mannhemia haemolytica* was detected in 17/299 nasal swabs (5.7%). Also, the carriage density of *H. somni* ranged between 10-100 genome copies/ml in the majority of the swabs (82.5%). For *M. haemolytica* the carriage density ranged between 100 and 1,000,000 genome copies/ml and *P. multocida* ranged between 1,000-100,000 genome copies/ml. Also, Kishimoto et al. (2017) developed a one-run qPCR to detect the BRD-bacterial pathogens from nasal samples taken of 40 animals collected at six different farms. This study targeted the bacterial genes *sod*A*, kmt*-1, and 16S rRNA and identified *M. haemolytica* in 12 samples, *P. multocida* in 23 samples, *H. somni* in 18, and *M. bovis* in 20 samples. Nevertheless, the carriage or the total number of copies of each bacterium in the nasal swabs were not calculated. In a different study, quantification of BRD associated pathogens was perform using the bacterial cells from tracheal aspirate samples pure culture collected from healthy and cattle with pneumonia (Pansri et al., 2020). For *M. haemolytica*, the bacterial load was between 0.4-4.7 $\log_{10}$ CFU/0.5 mL, for *P. multocida* was 0.5-5.9 $\log_{10}$ CFU/0.5 mL, for *H. somni* the density was 1.1-4.8 $\log_{10}$ CFU/0.5 mL, and for *M. bovis* the density was 1.1-3.3 $\log_{10}$ CFU/0.5 mL. Nonetheless, one pitfall of this study is that no comparison in the BRD-associated bacterial load were compared between the samples collected from healthy and cattle with pneumonia.

Even though qPCR can be useful in detecting and measuring the bacterial load of specific species in a sample, it is also important to determine the relative abundance of the desired bacteria in the

whole microbial community. By analyzing the relative abundance and number of copies per sample of the BRD-associated pathogen within an environment, it can provide more information on how these bacteria might interact with other non-BRD associated microbes in the cattle respiratory tract. It can also be an indicator of how the abundance of BRD-associated bacteria might be different from healthy and sick animals, a parameter that can be used to diagnose cattle with BRD.

2.6.5 Bacterial 16S rRNA gene PCR amplicon sequencing.

Next-generation sequencing of 16S rRNA gene PCR amplicons has been the most recently used method to characterize the cattle respiratory tract microbial community. The advantage of using this method is that it is possible to identify the bacteria present in the sample without relying upon in microbial culture. Additionally, 16S rRNA gene sequencing allows for the determination of microbial community diversity and relative abundance of individual taxa, while also allowing for the parallel sequencing of multiple samples at the same time (Gupta et al., 2019; Sontakke et al., 2009). Once the amplicons have been sequenced, the taxonomical classification of each sequence can be assigned using nearest-neighbor or naïve Bayesian classifier (Mizrahi-Man et al., 2013) algorithms.

Because microbial community diversity can be determined using 16S rRNA gene amplification, different studies identified that BRD animals had lower alpha diversity (that indicates the richness and evenness of the community) in the upper and lower respiratory tract than their healthy pen-mates (Holman et al., 2015a; McMullen et al., 2019; Timsit et al., 2018). These results elucidate the status of the bacterial community in the animal's respiratory microbiome during BRD. Because it is believed that a more diverse and stable microbial community is more likely to resist colonization from pathogens, and low diversity commonly causes low stability and reduced functional diversity (Cardinale et al., 2012; Dunlap, 2001). Thus, measuring diversity could be used as an indicator of sickness and could lead to a better understanding of how the disease develops. Different studies using 16S rRNA sequencing method in DNA extracted from nasal and nasopharyngeal swabs, had identified that cattle diagnosed with BRD, presented a lower bacterial richness and evenness that healthy animals (Zeineldin et al., 2020b; Timsit et al., 2018; Holman et al., 2015a) and it's believed that low microbial diversity could lead to pathogen colonization

(Cardinale et al., 2012; Knapp et al., 2017). These results could explain another animal susceptibility to BRD.

The 16S rRNA gene amplicon sequencing helped identify the commensal microbial community found in the respiratory tract of healthy cattle. At a phyla level, Proteobacteria, Firmicutes, Bacteroidetes, Tenericutes, Actinobacteria, and Fusobacteria are the most common; whereas at the genus level, the most common bacteria are *Mycoplasma, Acinetobacter, Filobacterium, Psychrobacter,* and *Moraxella* (Holman et al., 2017; Zeineldin et al., 2020a). Nonetheless, changes in specific taxa can be observed in different locations in the respiratory tract. The microbial community present in the cattle's hard palate, floor of the mouth and oropharynx has been shown to be mostly composed of bacterial genera *Streptococcus*. Then, the bacterial community in the cattle's nostrils and nasopharynx are mostly composed of *Moraxella* and *Mycoplasma*. The bacterial community on the tonsils are mainly composed of *Fusobacterium*, whiled the trachea and lung microbiome is composed mostly of bacteria from the *Mycoplasma* genera (McMullen et al., 2020). Thus, the 16S rRNA gene sequencing provides information of the commensal respiratory microbiome in cattle, information that could be used to identified microbes that are not commensal but pathogenic and that might be associated in the BRD development.

It is believed that cattle experience BRD due to an increase of pathogenic bacteria in the respiratory tract (Holman et al., 2015a). Thus, the use of 16S rRNA gene sequencing has been able to determine the bacteria present in sick from healthy animals. Lima et al. (2016) identified that the relative abundance of the bacteria genera *Mannheimia, Moraxella* and *Mycoplasma* were significantly higher in the animals with pneumonia than healthy animals. In a different study comparing the nasopharyngeal microbiota community of BRD and healthy animals 0 and 60 days after feedlot entry, BRD animals on day 0 had either *Mannheimia haemolytica* and *Pasteurella multocida;* however, after treating the animals for BRD, no BRD species were detected in day 60 (Holman et al., 2015a). As previously described, differences in the microbiome composition could depend on the sampling site in the animal. A study compared the upper and lower respiratory microbiome between healthy and BRD animals (Zeineldin et al., 2020b). In the upper respiratory microbiome, BRD animals presented significant higher predominance of *Acinetobacter, Mannhemia, Psychrobacter, Flavobacterium* and *Solibacillus* than the healthy animals. In the

lower respiratory microbiome, the genera *Mannheimia* and *Pasteurella* were significantly increased in the BRD animals than healthy. Nonetheless, presence of BRD-associated bacteria is sometimes not identified as differentially abundant between BRD and healthy cattle. Zeineldin et al. (2017) collected nasopharyngeal samples from healthy and BRD-affected cattle. In this study, BRD animals had many differentially abundant taxa compared with healthy animals. *Proteobacteria, Pseudomonadales, Moraxellaceae, Acinetobacter* were the best bacterial biomarkers of BRD-affected cattle; however, no BRD-associated pathogen were identified as biomarkers of the BRD group. In conclusion, it is possible to conclude that next-generation sequencing using 16S rRNA gene amplicon sequencing is a method that can be used to detect and identify differential bacterial abundance in the cattle's respiratory tract between BRD affected and healthy cattle.

2.7 Relations between the presence of BRD-bacteria and commensal bacteria

Nasopharyngeal microbiome plays an essential role in the animal's health status. When the animal is under stress, opportunistic bacteria found in the nasopharyngeal tract can translocate to the host cells, causing BRD (Amat et al., 2019a; Mosier, 2014). Also, there is a hypothesis that the resident microbial community in the respiratory tract might enhance or prevent BRD-associated bacteria from increasing in their abundance, causing an infection (Corbeil et al., 1985; Timsit et al., 2018). Probiotic bacteria strains like *Lactobacillus* strains, *Lactococcus locus*, *Paenibacillus polymixa* were shown to have had higher adhesion to the bovine bronchial epithelial (BBE) cells than the bacteria genera *Mannheimia haemolytica*, as well as the genera *Lactobacillus*, *Streptococcus*, and *Enterococcu*s were able to inhibit the growth of *Mannheimia haemolytica* when using culture-based methods (Amat et al., 2017, 2019b; Timsit et al., 2020) Negative correlations had been detected between lactic-acid producing bacteria with the Pasteurellaceae family in the nasopharyngeal microbiome of cattle transported to the feedlot (Amat et al., 2019a). In a different study, the nasopharyngeal microbiota of healthy and animals with bronchopneumonia (BP) were compared across four different feedlots (Timsit et al., 2018). Animals with BP had *Mycoplasma bovis, Mannheimia haemolytica,* and *Pasteurella multocida*, while healthy animals had mostly *Mycoplasma dispar, Lactococcus lactis,* and *Lactobacillus casei*. These results might give an insight into the bacterial competition occurring in the cattle's

respiratory microbiota and could be used to understand the reason behind why some cattle do or do not develop BRD.

2.8 Relationship of the nasal microbiome to the lower respiratory tract

BRD development has been linked to the presence of pathogens in the respiratory tract. Thus, intensive research had been developed to identify these and characterize the cattle's respiratory microbiome. Anatomically speaking, the cattle respiratory tract can be divided into two sections: upper respiratory tract (URT) that includes nasal cavities, paranasal sinuses, nasal passages, nasopharynx, oropharynx, tonsils, and the upper portion of the larynx and the lower respiratory tract (LRT) that includes the lower portion of the larynx, trachea, bronchi, bronchioles, and alveoli (Beers, 1999). The process of characterizing the LRT in cattle has been dependent on culture-based methods of lung tissue and only pathogens that can be cultured had been identified (Griffin et al., 2010; Fulton & Confer, 2012); on the contrary, URT is more easily sampled. The most common method is through the use of nasal or nasopharyngeal swabs, and these approaches are less invasive and take less time than collecting lung tissue (Pardon & Buczinski, 2020). Studies had identified correlations in specific taxa between the URT and LRT that give a sense of mutualistic interrelationship between the two microbial communities (Zeineldin et al., 2017). Also, Zeineldin et al., (2020b) identified 50 taxa to possess strong correlations between their presence in the URT and LRT. Doyle et al., (2017) identified agreement in BRD pathogens isolated regardless of the sample type collected: nasal, nasopharyngeal, bronchoalveolar lavage and transtracheal wash . Thus, these results bring supportive information that characterizing the nasal microbiome could be used to predict the microbiome in the LRT; however, more research is needed in this area. Unfortunately, it is important to emphasize that nasal swabs are susceptible to be contaminated; hence, it is necessary to clean the cattle's nasal cavity before sampling and use negative controls (clean swabs) as a way to identify any contaminant present in the swabs (Fulton & Confer, 2012; Zeineldin et al., 2019).

2.9 Summary

In recent years, intensive research has been developed to diagnose cattle with BRD from healthy animals. One common method utilized is molecular techniques (e.g., PCR, qPCR, 16S rRNA gene

sequencing) to identify BRD-associated bacteria in the respiratory tract of healthy and BRD- animals. Even though these methods could detect and measure BRD-associated bacteria between the two groups, a gap in these studies is the interaction and relation of BRD bacteria with other microbes, such as commensal bacteria, viruses, and fungus, have not been taken into consideration. So far, studies were able to measure the BRD bacterial load in the cattle respiratory tract; nevertheless, no threshold of the bacterial load has been established to diagnosed BRD animals. We hope that with the finding from this study, we will identify an interaction of BRD-associated pathogen within the cattle respiratory tract and what other interactions might occur with commensal bacteria. To determine if a bacterial load threshold exists between BRD-associated bacteria present in healthy and BRD animals to use as BRD diagnosis. Finally, it is necessary to understand that Bovine Respiratory Disease can be caused by multiple factors (e.g., predisposing, environmental, and epidemiological). Thus, the analysis of these factors and their interactions with the respiratory microbiome can bring more information on when the disease begins and how it can be treated efficiently.

CHAPTER 3. MATERIALS AND METHODS

3.1 Cattle nasal swab collection

3.1.1 Animal population and selection

A total of 133 Holstein steers of approximately 6-7 months old, housed in the same environment were sampled in our study from July to December 2020 at the Ault Farm located in Indiana, US. Animals with BRD were identified following the DART approach that focuses in the clinical signs as depression, appetite loss, respiratory character change, and increased rectal temperature (Griffin et al., 2010). We will refer to animals identified to have BRD according to the DART method as BRD animals. Once an animal was identified as positive for BRD, healthy pen mates were also selected for nasal swab sampling. The animals selected for the study could not had been previously treated for BRD with antibiotics (individually or mass-medicated) or treated for any disorder with antibiotics for the previous 100 days. A total of 75 healthy and 58 BRD-positive animals were included in the study. Animal records were checked 3 months after sampling and if any animal that initially was considered a health animal was treated with antibiotics, it was removed from the study.

3.1.2 Cattle nasal swab collection

After identifying the healthy and BRD-animals, two nasal swabs were collected per animal. Before sampling, the nostrils were cleaned with wipes to remove any nasal discharge and dirt. Two unguarded swabs were inserted simultaneously about 3-5 cm deep into the right nostril and then into the left nostril. Nasal swabs were inserted into an empty tube labelled with the animal ID and date and then, transported to the lab on ice for processing. During sampling, rectal temperature, prescribed treatment, date of sample, and pen ID were collected per animal for further data analysis.

3.2 Nasal swab DNA extraction

Nasal swabs collected from healthy, and BRD-animals were transported to the lab. The bacterial and mucosal content was extracted from the swab's prior DNA extraction. The tip of the swab was cut and inserted into new tube and 1 mL of Nuclease-free water was added. Then, the tubes

containing the swab's tip were vortexed horizontally for 5 min, the swab tip was discarded, and the tubes were centrifuge 6000 x *g* for 10 min, this allowed the separation of the supernatant and the pellet. The pellets and supernatant were stored at -20°C and -80°C respectively until further processing.

Total DNA was extracted from the pellets using the DNeasy Blood & Tissue Kit (Qiagen, Germantown, MD, USA) following Holman et al., (2015) protocol. The extraction process combines physical, chemical and enzymatic lysis methods to extract the DNA from the bacterial cells. In the protocol, the pellets were re-suspended in 180 µl of enzymatic lysis buffer containing the enzymes mutanolysin (300 U mL^{-1}) and lysozyme (20 mg ml^{-1}). The mixture was vortexed for 20 s and incubated at 37°C for 1 h in a thermomixer at 300 rpm. Proteinase K (25 µl) and 200 µl of Buffer AL (provided in the kit) were added; the mixture was vortexed and incubated at 56°C for 30 min. After incubation, the mixture was transferred into a PowerBead Tubes containing 0.1 mm glass beads (Qiagen, Germantown, MD, USA, not provided in the kit) and mixed using the TissueLyser II (Qiagen) at 30 Hz for 5 min, followed by centrifugation at 13,000 x *g* for 5 min. The supernatant was transferred into a new tube and 200 µl of ethanol were added to the mixture. From this point on, the DNeasy Blood & Tissue Kit was used following the manufacturer's instructions. After extraction, the DNA quality was measured using Nanodrop 2000/2000c Spectrophotometer and the concentration was measured with Qubit 4 (Thermo Fisher Scientific, PA, USA). DNA was stored at -20°C until further processing. An empty tube was extracted and sequenced along with the samples to determinate the presence of contaminants (DNA extraction and sequencing negative control).

3.3 16S rRNA gene amplicon library preparation and sequencing

The extracted DNA samples from the animal and empty tubes, were used to create a 16S rRNA gene amplicon library. This library was constructed using a barcode indexed amplification product from the V4 region of the 16S rRNA gene following the protocol described by Kozich et al., 2013. PCR amplification was performed using AccuPrime™ Pfx SuperMix (Thermo Fisher Scientific, MA, USA) following Kozich et al., 2013 protocol. PCR-grade water was used as negative control and a mock community (20 Strain Even Mix 138 Genomic Material; ATCC® MSA-1002TM) as positive control. PCR amplification quality was checked via gel electrophoresis. Amplified DNA

was normalized using SequalPrep Normalization Kit, and 5 µl of the amplified DNA was used to create a pool at equal concentration. The amplicons were sequences via Illumina MiSeq Sequences (2x250 paired-end) at the Purdue Genomic Core Facility.

3.4 Bioinformatics analysis

3.4.1 16S rRNA gene library analysis

Raw sequences data obtained from the 16S rRNA gene fragment's amplification were analyzed using Quantitative Insight Into Microbial Ecology (QIIME2) v.2020.2. With QIIME2, it is possible to identify how many sequences were obtained per sample. In the study, using the DADA2 (Callahan et al., 2016) denoising step, the forward and reverse sequences were trimmed at position 0, and the forward and reverse sequences were truncated at position 251 and 223, respectively, to obtained sequences with a quality > Q30. A total of 15,287,698 sequences were identified before the denoising step (DADA2) and 13,127,373 sequences after denoising. All the sequences were clustered into Amplicon Sequence Variants (ASVs) with 100% similarity.

To calculate the alpha and beta microbial diversity in the nasal swab samples, the ASV table was rarefied to 40,420 sequences per sample. In this step, nine samples were lost due to low sequence count. Alpha diversity measures the richness and evenness of an environment. Richness is an indicator of how many different species are present in the sample whereas, evenness is an indicator of the abundance distribution of species within each sample (Hagerty et al., 2020; Whittaker, 1960). Alpha diversity was estimated in QIIME2 using the Observed OTUs and Chao1 metrics for richness, Pielou index as a measure of evenness and Faith phylogenetic diversity (Faith_Pd) was also calculated (Chao, 1984; DeSantis et al., 2006; Faith, 1992; Pielou, 1966).

The beta-diversity estimates the pairwise microbial community structure dissimilarity. Beta diversity between animals were analyzed using Bray-Curtis Dissimilarity Index and Weighted UniFrac (incorporates phylogenetic relation) (Lozupone & Knight, 2005) methods and plotted as principal coordinate analysis (PCoA) using RStudio (v1.3.1093). To test the difference in beta diversity, a permutational multivariate analysis of variance test (PERMANOVA; $P \leq 0.05$) using the vegan package (Oksanen et al., 2019). In addition, a dispersion test was performed to determine

the distance of the samples of the two groups (BRD or healthy animals) using the vegan package (Oksanen et al., 2019), followed by a permutation test of multivariate homogeneity of groups dispersion using the vegan package (Oksanen et al., 2019).

The taxonomy was assigned using SILVA 13_8, 515F/806 region database. A negative binomial distribution method, DESeq, was used to determined differentially abundant taxa between BRD and healthy animals. The test was performed using the package DESeq2 (Anders & Huber, 2010) in RStudio. ASVs with a log2 fold change > 2 and statistical significance of $P \leq 0.05$ were selected as differentially abundant ASV between BRD and healthy animals.

3.4.2 Mock Community and Empty Swab sequencing analysis

Mock community analysis

A mock community (ATCC® MSA-1002TM) was used as a positive control for the amplification and sequencing step. To determine the amplification and sequencing quality, we compared the sequences obtained from the mock community with a reference containing the correct sequences for the 20 known strains. The reference file contains the 16S rRNA gene sequences of all 16S rRNA copies from the 20 strains present in the mock community sample.

The raw sequence data was analyzed using Quantitative Insight Into Microbial Ecology (QIIME2) v.2020.2. The raw reads were trimmed using DADA2 to remove any sections with a low-quality score (< Q30). To evaluate the sequencing quality, we used the Qiime2 (v.2-2020.2) function evaluate_seqs, which aligns the observed sequences to the set of expected or reference sequences, and determined the matches and mismatches (Camacho et al., 2009).

Empty tubes DNA sequencing analysis.

Raw sequence data obtained from the empty tubes used as a negative control during DNA extraction and sequencing were also analyzed using QIIME2 following the procedure explained above. Forward and reverse sequences were trimmed at 0, and truncated at positions 250 and 220 respectively. The taxonomy was assigned using SILVA 13_8, 515F/806 region database.

Presence of contaminants during DNA extraction and sequencing.

DNA contamination can arise from many sources such as sample collection, laboratory equipment, DNA extraction kits (reagents), amplification and sequencing reagents (Nguyen et al., 2015). Because of this, an empty tube was extracted and sequenced along with the samples. This helped determine if contamination occurred during the process. To determine the ASVs that are considered contaminants, the mock community (positive control) taxonomical composition was compared to the mock community reference that contains 32 known bacterial strains. Any ASVs in the mock community (positive control) that did not match mock community reference were considered contamination. Then, the contaminants were compared to the ASVs present in the empty tubes; any ASVs shared between the two groups were considered contaminants. Lastly, the taxonomical composition of the empty tube was compared to the samples; any ASVs matched to the samples and when combined, present a relative abundance >10% of the community in the empty tubes were considered contaminants and were removed from the data.

3.5 qPCR analysis

Gene amplicon generation for BRD-pathogens and qPCR standard curve generation

DNA extracted from pure isolates of *Pasteurella multocida, Histophilus somni,* and *Mannheimia haemolytica* acquired from the Indiana Animal Disease Diagnostic Laboratory (ADDL) at Purdue University and DNA from *Mycoplasma bovis* strain 25523 (ATCC) was used to generate the qPCR standard curve and as bacterial positive control. PCR assays were performed to target *sod*A for *M. haemolytica*, 16S rRNA gene for *H. somni* and *P. multocida* and the gene *opp*D for *M. bovis* using previously published primers and fluorescent probes (Table 1) (Sachse et al., 2010; Thomas et al., 2019). Amplicon DNA concentration was determined by Qubit 4 (Thermo Fisher Scientific, PA, USA) and number of copies was calculated using the concentration and length of the amplicon.

The PCR assays for *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica,* and *Mycoplasma bovis* were performed in a 50 µl volume consisting of 25 µl of iTaq™ Universal Probes Supermix (BioRad, CA, USA), 12.5 µl Primer/Probe mix (IDT) listed above, 10 µl Nuclease-free water and 2.5 µl of nucleic acid template. The concentrations for the primers and probes for the four bacteria were 300 nM and 100 nM as reported in Thomas et al. (2019). PCR

assays were performed in an Eppendorf Mastercycler Gradient Model 533 and the cycling conditions for *H. somni* and *P. multocida* were 95°C for 3 min followed, by 40 cycles of 95°C for 3 s and 60°C for 60 s. Cycling conditions for *M. haemolytica* were as follows: 95°C for 3 min, followed by 40 cycles of 95°C for 3 s and 69°C for 60 s. Cycling conditions for *Mycoplasma bovis* were 95°C for 10 min, followed by 45 cycles of 95°C for 30 s, and 60°C for 60 s.

Table 1. BRD-associated bacteria primers and probes sequences for PCR and qPCR reactions.

Target	Target gene	Primer name	Sequence (5'-3')	Size (bp)	Ref
M. haemolytica	*sod*A	Mh-SGF	AGCAGCGACTACTCGTGTTGGTTCAG	26	1
M. haemolytica	*sod*A	Mh-SGR	AAGACTAAAATCGGATAGCCTGAAACGCCTG	31	1
M. haemolytica	*sod*A	Mh-BV1P*	TTCAACCGCTAACCAGGACAACCCAC	26	1
P. multocida	16S rRNA	Pm-TMF	CGCAGGCAATGAATTCTCTTC	21	2
P. multocida	16S rRNA	Pm-TMR	GGCGCTCTTCAGCTGTTTTT	20	2
P. multocida	16S rRNA	Pm-TMP*	ACTGCACCAACAAATGCTTGCTGAGTTAGC	30	2
H. somni	16S rRNA	Hs-TMF	AGGAAGGCGATTAGTTTAAGAGATTAATT	29	2
H. somni	16S rRNA	Hs-TMR	TCACACCTCACTTAAGTCACCACCT	25	2
H. somni	16S rRNA	Hs-TMP*	ATTGACGATAATCACAGAAGAAGCACCGGC	30	2
M. bovis	*opp*D	PMB996-F	TCAAGGAACCCCACCAGAT	19	3
M. bovis	*opp*D	PMB1066-R	AGGCAAAGTCATTTCTAGGTGCAA	24	3
M. bovis	*opp*D	Mbovis1016*	TGGCAAACTTACCTATCGGTGACCCT	26	3

1Guenther et al., 2008
2 Mahony et al., 2007
3 Sachse et al., 2010
** Fluorescence probes*

PCR-grade water was used as negative control each PCR assay. PCR amplification quality was checked via gel electrophoresis. *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica* and *Mycoplasma bovis* gene amplicons were cleaned and purified using Monarch PCR & DNA Cleanup kit (New England BioLabs, MA, USA). Purified gene amplicons were stored at -20°C until further use.

For the qPCR standard curve generation, 10-fold serial dilutions (10^8 to 10^0) were created for each BRD-associated bacterial amplicon. The qPCR technical triplicate assays were performed in 20 µl total volume containing 10 µl iTaq™ Universal Probes Supermix (BioRad, CA, USA), 5 µl Primers/Probes and 5 µl of each BRD-associated bacteria amplicon. The qPCR assays were performed in CFX96 Real-Time System Thermal Cycler (BioRad, CA, USA). *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica* and *Mycoplasma bovis* cycling conditions

as well as the primers and probe concentration were the same as the BRD-associated bacteria PCR conditions as described above.

The standard curve was generated using a linear regression of technical triplicate average cycle quantification (Cq) and $\log_{10}$ amplicon copies/µl as known from each dilution. Amplification efficiency (*E)* was calculated using the slope of the standard curve and the formula: *E* (%) = (-1+10^(-1/slope)) x 100. When performing the qPCR assays for each of the bacterium and 16S rRNA, one dilution used to create the standard curve was used as the positive control and PCR-grade water as the negative control in each of the replicates per qPCR assay.

Gene amplicon generation for 16S rRNA and qPCR standard curve generation

Because the 16S rRNA gene is present in all bacteria, qPCR targeting the 16S rRNA gene was performed to quantify the abundance of the total bacterial microbial community in the nasal samples from the healthy and BRD-positive animals. DNA extracted from the cattle nasal samples was used as the nucleic acid template in the PCR reaction. 16S rRNA gene PCR assay was performed in a Eppendorf Mastercycler Gradient Model 533, using the bacteria specific primer 8F (5' AGAGTTTGATCCTGGCTCAG 3') and universal primer 1492R (5' ACGGTTACCTTGTTACGACTT 3') to obtain the bacterial 16S rRNA gene amplicon that was used as the qPCR nucleic acid template (Galkiewicz & Kellogg, 2008). PCR assays were performed in 50 µl volume reaction consisting of 42.5 µl of AccuPrime™ Pfx SuperMix (Thermo Fisher Scientific, MA, USA), 2.5 µl of each primer (8F/1492R), 1.5 µl of Nuclease-free water and 1 µl DNA template. The primer concentration and PCR cycling conditions were performed as stated in Kozich et al., 2013 protocol. PCR-grade water was used as negative control and a mock community (20 Strain Even Mix 138 Genomic Material; ATCC® MSA-1002TM) as positive control. PCR amplification quality was checked via gel electrophoresis. 16S rRNA gene amplicons were cleaned and purified using Monarch PCR & DNA Cleanup kit (New England BioLabs, MA, USA). Purified 16S rRNA gene amplicons were stored at -20°C until next step.

For the qPCR standard curve generation, 9-fold serial dilutions (10^8 to 10^0) were performed with the 16S rRNA gene amplicons. The qPCR assays were performed in 20 µl total volume containing 10 µl LightCycler 480 SYBR Green I Master (Thermo Fisher Scientific, PA, USA), 5 µl

Primers/Probes and 5 µl of the gene amplicons. The qPCR assays were performed in a CFX96 Real-Time System Thermal Cycler (BioRad, CA, USA) using the universal bacteria primers 1132F and 1108R (Leigh et al., 2007). Each primers concentration was 6 pmol and the cycling conditions for qPCR were 40 cycles of 95°C for 15s and 60°C for 1 min as stated in Leigh et al. (2007). The standard curve was generated using a linear regression of cycle quantification (Cq) versus $\log_{10}$ #amplicon copies/µl value obtained from each dilution. Three technical replicates were performed at each dilution to obtain the Average Cq value needed for the standard curve generation. Amplification efficiency (*E)* was calculated using the slope of the standard curve and the formula: *E*(%) = (-1+10^(-1/slope)) x 100. In each replicate, a standard dilution was used as the positive control and PCR-grade water as the negative control.

BRD-pathogens and relative abundance.

The prevalence of *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica* and *Mycoplasma bovis* were calculated from the animals that tested positive in each qPCR assay from the total numbers of samples. To consider an animal positive to the bacterium, the Cq value corresponding to the endpoint of the standard curve was used as a threshold (Table 2). The prevalence of the BRD-pathogens was also calculated for the healthy and BRD-animals. The relative abundance of the four bacteria was calculated by dividing the total number of gene copies for each bacterium by the sum of *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica* and *Mycoplasma bovis* gene copies in the samples.

Table 2. Evaluation of *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica* and *Mycoplasma bovis* qPCR assay using nucleic acids templates generated form PCR reactions.

qPCR assay		***Pasteurella multocida***	***Mannheimia haemolytica***	***Mycoplasma bovis***	***Histophilus somni***	**16S rRNA gene**
Standard equation	Slope	-3.300	-3.145	-3.343	-3.596	-3.011
	Intercept	36.921	35.920	37.098	39.723	35.770
	Replicates	3	3	3	3	3
Efficiency (%)		100.94	107.96	99.13	89.71	114.84
Dilutions		9	9	9	9	9
Cq cut-off value$^{\Psi}$		33.53	31.76	32.1	36.77	30.52

$^{\Psi}$ Cq value corresponding to the last dilution in the standard curve (10^0) at which samples tested positive.

3.6 Linear Discriminant Analysis to classify BRD and Healthy animals

As previously described, BRD is caused by multiple factors, making the process of BRD diagnosis a challenge to producers. This study aims to identify a possible combination of BRD-associated pathogen copy number and any other factor that could help discriminate BRD-positive animals from healthy animals. A linear discriminant analysis using the MASS package (Ripley et al., 2021) in RStudio was performed to identify a possible combination between BRD-associated bacteria gene copy number, animal age, rectal temperature, and 16S rRNA gene copy number that could be used to classify BRD-positive and healthy animals. To determine the best model combination, the sensitivity (Equation 1), specificity (Equation 2), percentage of true positives that indicates the agreement between the LDA model and visualization of clinical signs in classifying BRD animals, and true negatives that indicates the agreement between the LDA and visualization of clinical signs in classifying healthy animals, and misclassification rate were evaluated for each of the combinations. The best possible model was selected using the following criteria: high sensitivity, specificity, true positive percent (TP%) that indicates the true positive animals identified by LDA model when using the visual observation as the reference, true negatives percent (TN%) that indicates the true negative animals identified by LDA model when using the visual observation as the reference, and low misclassification rate.

Two types of Linear Discriminant Analysis were performed. The first analysis included all available samples. Three samples were not included due to missing rectal temperature data, the LDA model could not be run if there is missing data. The second analysis only included the samples that tested positive to *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica,* and *Mycoplasma bovis.*

$$Sensitivity\ (Sen) = \frac{\#\ of\ true\ positives}{\#\ of\ true\ positives + \#\ of\ false\ negatives} \quad (1)$$

$$Specificity\ (Spec) = \frac{\#\ of\ True\ negatives}{\#\ of\ True\ negatices + \#\ of\ false\ positives} \quad (2)$$

3.7 Statistical analysis for 16S rRNA gene sequencing and qPCR data

Alpha diversity metrics (Observed ASVs, Chao1, Pielou, and Faith_Pd) and BRD-associated pathogen's gene copy number and relative abundance were analyzed using a General Linear Mixed Model with the random factor specified to only include random slopes using the afex package

(Singmann et al., 2021) in RStudio. Health status (BRD or Healthy) was included as a fixed factor and Pen as a random factor. The age and date of sampling were included as continuous factors in the model. Assumptions of normality of the residuals and homogeneity of variance were checked using the afex package. Dependent variables that did not meet the assumptions were log-transformed. Statistical significance was defined as $P \leq 0.05$.

All the statistical involving the microbiome composition were performed in RStudio (v. 1.3.1093. The codes and metadata are included in the Appendix B section of this study. Additional files used in this study are available at https://github.com/EuniceCenteno/BRDNasal for reference and reproducibility.

3.8 Microbial community co-occurrence analysis

In addition to the taxonomical analysis, a co-occurrence analysis was performed to identify ASV pairs present and most prevalent in the BRD and healthy microbial community and the ASV pairs that are less likely to co-exist in the samples. The table containing all the observed ASVs identified in the samples was subset into two tables: samples from BRD-positive and healthy animals. ASVs with an average relative abundance < 0.0001 were removed from the dataset, and the reads count were rarified to 40,420 reads per sample; this process removed rare ASVs that contributed noise to the dataset. A total of 1236 ASVs were used in the analysis. To perform the co-occurrence analysis, the ASV tables were converted to a binary format: presence or absence. Any ASV with an abundance greater than 0 was converted to 1, indicating that the ASV was present in the sample; any ASV with an abundance of 0 remained 0, indicating absence. Co-occurrence analysis was performed with the package cooccur (Griffith et al., 2016) in RStudio (v1.3.1093) with R (v. 4.0.3). This function identifies species pairs with positive, and negative associations, and random associations. Positive association indicates the AVS that are more likely to co-occurred in more locations greater that what would be expected if they were randomly distributed relative to the other one. Negative associations indicated AVSs pair combinations that co-occurred in fewer location that what was expected. Random that indicates two species are distributed randomly or independent from one another, and it identifies in how many sites (samples) the pairs were observed. Once the species pairs with positive association, were identified for the BRD and healthy dataset, a set of rules were followed to determine how common it is to observe the pairs in the groups:

1. ASV pairs must be present in at least 60% of the total number of samples: BRD n > 34 samples and Healthy n > 45 samples.

2. The probability that the two species will occur in the sample greater than 0.9.

With the negative associations, a different set of rules were applied to identify the ASVs pair combination that were less likely to co-occurred in the BRD and healthy groups:

1. Probability that the two species would co-occur at a frequency less than the observed number of co-occurrence sites if the two species were distributed randomly (independently) of one another lower than 0.01.

2. The probability that the two species will occur in the sample lower than 0.05.
3. ASVs pair combinations observed co-occurrence must be lower than <1 samples.

Once the species pairs combinations were identified, any pairs combination with the BRD-pathogens were selected. This allowed us to see the level of association of the pathogens with any other bacteria present in the nasal cavity of healthy and BRD-positive animals.

CHAPTER 4. RESULTS

4.1 16S rRNA gene sequencing analysis

4.1.1 Mock community and empty tubes analysis

Three mock communities were amplified and sequenced, along with the DNA extracted from the cattle's nasal swabs. A total of 54 amplicon sequence variants (ASVs) were identified in the three mock communities used in the study, with 648,634 sequences. After comparing the unknown ASVs to the 20 known bacteria strains, 18 out of 20 of the reference bacteria were identified, with an E-value between 3.93 E-33 and 1.23 E-135. A total of 758 mismatches and 87 gaps were found between the query and the 20 known strains. For the empty tubes, a total of 719 ASVs and 2,706,853 sequences were observed in the study. Based on the taxonomical identity, the combined relative abundance of *Pseudoalteromonas* and *Vibrio* composed >50% of community in empty tubes (Fig. A1).

4.1.2 Contaminant during DNA extraction and sequencing.

The mock community's taxonomical composition was compared to a mock community reference file containing 20 known bacterial strains. From the 54 ASVs observed in the mock community, 50 matched the mock reference, and 4 ASVs did not match the reference. The mock contaminant ASVs were assigned to the class Chloroplast, order *Ruminococcaceae*, and species *Lysobacter enzymogenes.* No shared ASVs were observed between the mock contaminants ASVs and empty tubes taxonomical composition. Only one ASV belonging to the order Chloroplast was shared between the mock contaminants and nasal samples. After comparing the taxonomical composition of the empty tubes and the samples, a total of 299 ASVs were shared between the two groups. As mentioned above, the genera *Pseudoalteromonas* and *Vibrio* each presented a relative abundance greater than 10% in the empty swabs. Because of this, the two genera were considered contaminants and were removed from the samples.

4.1.3 Nasal microbiome alpha diversity

The 16S rRNA gene sequencing is one of the most common methods used to study the bacterial community and composition within a set of samples. In this study, DNA extracted from nasal swabs of 124 samples isolated from pen-mates' cattle were sequenced targeting the 16S rRNA V4 gene region. A total of 15,287,698 sequences were identified before the denoising step (DADA2) and 13,127,373 sequences after denoising. A total of 18,010 Amplicon Sequence Variants (ASVs) were observed in the study. After rarefying the total number of reads to 40,420 per sample, 16,376 different ASVs remained and were used to quantify the nasal alpha and beta diversity.

In this study, the richness predicted by Observed ASVs (F 1, 114.18 = 13.375, $P < 0.0001$; Fig. 4.1) and Chao 1 (F 1, 114.38= 12.0456, $P < 0.0001$, Fig. 4.1) significantly decreased in the BRD animals compared to healthy animals. Also, the evenness predicted by Pielou_e was significantly decreased in BRD animals compared to healthy animals (F 1, 113.67= 7.3700, $P < 0.007$; Fig. 4.1). In addition, the phylogenetic diversity predicted by Faith_pd was also significantly decreased in BRD animals compared to healthy animals (F 1, 112.82= 10.212, $P <0.001$; Fig. 4.1).

We also found in this study, that microbial richness measured with Observed ASVs (F 1, 120.00= 18.740, $P < 0.0001$; Fig. 4.2) and Chao 1 (F 1, 120.00=16.5301, $P < 0.0001$; Fig. 4.2), was decreased in nasal samples collected later in the time of the study. The same effect was observed for the microbial evenness (F 1, 119.99= 37.0722, $P < 0.0001$; Fig. 4.2) and phylogenetic diversity (F 1,119.89=15.047, $P< 0.0001$; Fig. 4.2), as both significantly decreased in nasal samples collected later in the time of the study. Lastly, the richness, Observed AVSs (F 1,119.87= 5.424, P <0.02) and Chao 1 (F 1,119.70= 4.9954, $P <0.02$), was significantly affected by the age of the animals. Also, age of the animals significantly affected the evenness (F 1, 119.99=6.6816, $P <0.01$) and phylogenetic diversity (F 1,119.34=4.007, $P<0.04$) of the nasal microbiome.

One possible factor that could had contributed to the decrease of the alpha diversity from the beginning to the end of the study is the environmental temperature. As described before, samples were collected from July (summer) to December (winter) 2020 in a farm located at Indiana, US. To identify if environmental temperature could have an effect on the difference in alpha diversity of the nasal cavity, average temperature of the dates when the samples were collected was retrieved

form WeatherUnderground.com. Average daily temperature data was retrieved from the closest weather station to the farm where the samples were collected. Correlation using Pearson's correlation indicated a significant correlation between the average temperature and date when the samples were collected in this study (t 122=12.901; R=0.76, P < 0.0001; Fig. A2). Indicating that the samples from the beginning of the study (July) were collected on days with high environmental temperature relative to the sampling days at of the study. Average daily temperature was tested to identify an effect in the cattle's nasal alpha diversity. A General Linear Mixed Model was performed using average daily temperature as a continuous factor and pen as random factor with random slope. The bacterial richness predicted by Observed ASVs (F 1, 110.82,=7.4379, P <0.007; Fig 4.3) and Chao1 (F 1, 111.12=6.4313, P < 0.01; Fig 4.3) , evenness predicted by Pielou_e (F 1, 112.5=13.595, P < 0.0004; Fig 4.3) and phylogenetic diversity predicted by Faith_pd (F 1, 115.68=5.4671, P < 0.021; Fig 4.3) were all significantly affected by the average temperature, indicated high bacterial richness when the environmental temperature is high (summer-July) and it decreases as the environmental temperature decreases (winter-December).

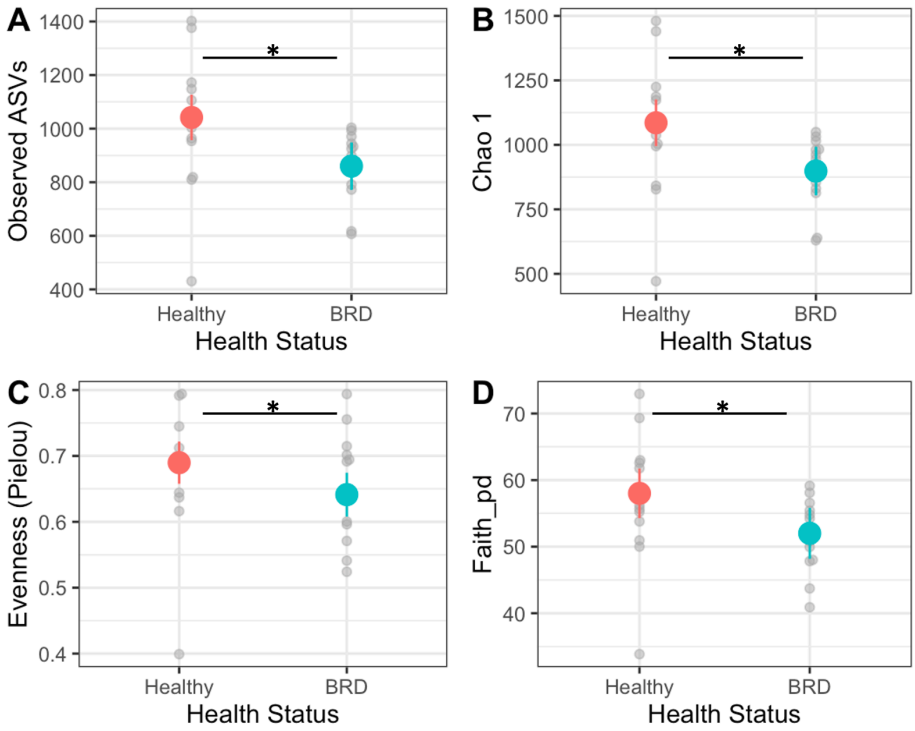

Figure 4.1. Cattle nasal alpha diversity between healthy and BRD animals. Observed ASVs and Chao 1 (A and B) measure the richness of the microbiome community. Evenness was measured with Pielou (C) and the phylogenetic relationship was measured with Faith (D). An asterisk (*) and horizontal line represent a statistical difference ($P \leq 0.05$) between the two groups. Colored circles represent the means of the BRD and healthy group and the gray dots represent the raw data of each group.

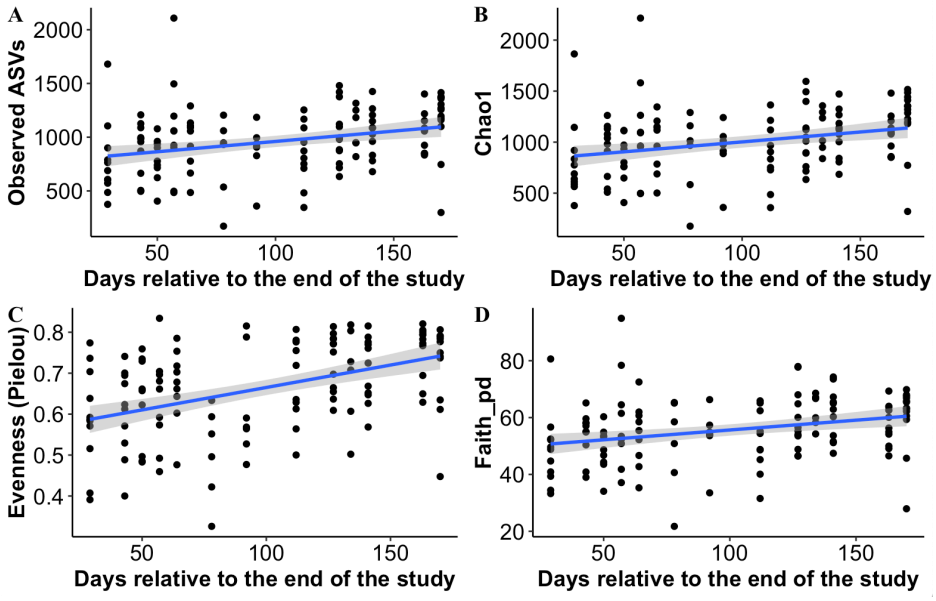

Figure 4.2. Variation in the cattle nasal alpha diversity predicted by Observed ASVs (A), Chao1 (B), Pielou (C) and Faith (D) (± 95% confident bands) relative to the date of collection. High values in the x axis represent the dates of samples collected at the beginning of the study (July, 2020); low values represent the dates of samples collected at the end of the study (December, 2020).

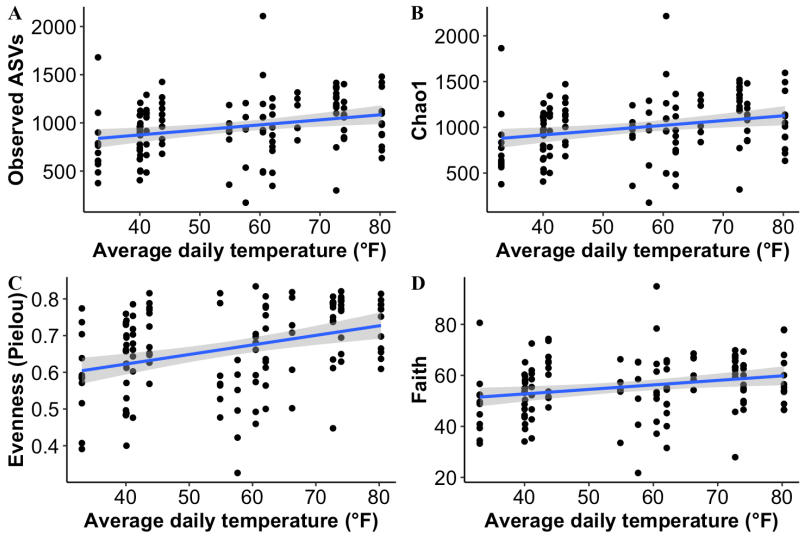

Figure 4.3. Variation in the cattle's nasal alpha diversity predicted by Observed ASVs (A), Chao1 (B), Pielou (C) and Faith (D) (± 95% confident bands) relative to the average daily temperature of when then samples were collected.

Nasal microbiome Beta Diversity

Nasal community structure (beta diversity) as determined by Bray-Curtis dissimilarity (F 1, 123= 2.1804, $P < 0.007$; Fig. 4.4) and Weighted UniFrac (F 1, 123= 2.002, $P < 0.03$; Fig. 4.4), was significantly different between BRD-positive and healthy animals. Beta diversity determines the community dissimilarity between two communities. Dissimilarities are often expressed as a distance between the two groups, and distance is directly proportional to dissimilarity. The distance between the group centroids was 0.0089 (Bray-Curtis) or 0.056 (Weighted UniFrac).

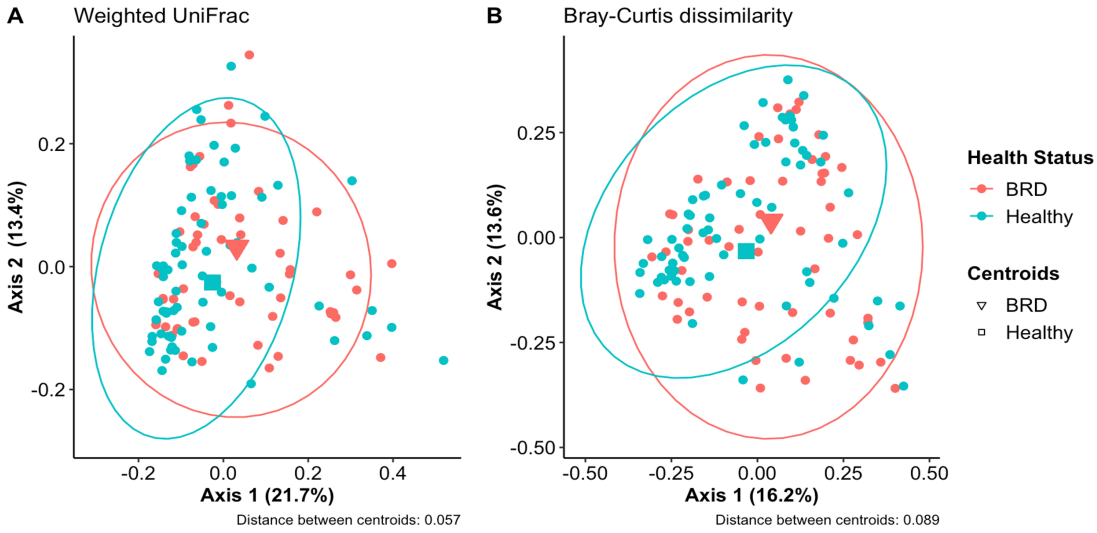

Figure 4.4. Principal Component Analysis (PCoA) determined by Weighted UniFrac distances (A) and Bray-Curtis Dissimilarity (B) between BRD and healthy animals. Ellipses indicate 95% confidence interval for each of the health status groups. Axis 1 represent the axis that explains the greatest amount of the variation followed by Axis 2. Shape indicates the centroids of the ellipses. Distances of the centroids between the two groups is indicated in the caption below each plot.

Cattle nasal microbiota composition and differentially abundant taxa between BRD and healthy animals

In this study, the top four most abundant phyla in the nasal microbiome from all animals were Proteobacteria (~30% of the community on average), Firmicutes (~20%), Bacteroidetes (~20%), and Actinobacteria (~10%) regardless of the health status (Fig. A3). At the family level, *Moraxellaceae* (22.23%), *Pasteurellaceae* (18.75%), *Corynebacteriacea* (9.68%) were the most abundant regardless of health status; followed by *Mycoplasmataceae* (3.54%) in BRD animals and *Weeksellaceae* (3.68%) in healthy animals (Fig. A4). Interestingly, at a genera

level, *Mannheimia* (5.19%), *Moraxella* (4.59%), and *Mycoplasma* (3.92%) were the most abundant in BRD animals, whereas healthy animals contained *Corynebacterium 1* (4.78%), *Moraxella* (4.48%), and *Mannheimia* (4.10%) (Fig. A5).

DESeq analysis was used to identify differentially abundant taxa in this study. A total of 15 ASVs were increased in BRD compared to healthy animals (log2 fold change >1, $P \leq 0.05$), and 8 ASVs were increased in healthy animals compared to BRD-positive animals (Fig 4.5). From the differentially abundant ASVs, the species *Mycoplasma alkalences* 14918 and *Mycoplasma arginini* had the highest log2 fold change, 3.081 and 1.987, in the BRD group. On the contrary, ASVs classified as unclassified *Moraxellaceae*, and uncultured *Gemmobacter* family members had the highest log2 fold change (2.807 and 1.529) in the healthy groups. Interestingly, in the BRD group, 4 ASVs were identified as members of the *Mycoplasma* genera. Only one *Mycoplasma* species (*Mycoplasma bovirhinis*) was increased in the healthy group; no BRD-associated pathogen species were identified as differentially abundant in the BRD or healthy group. However, no BRD-associated species presented a significant abundance between the two groups.

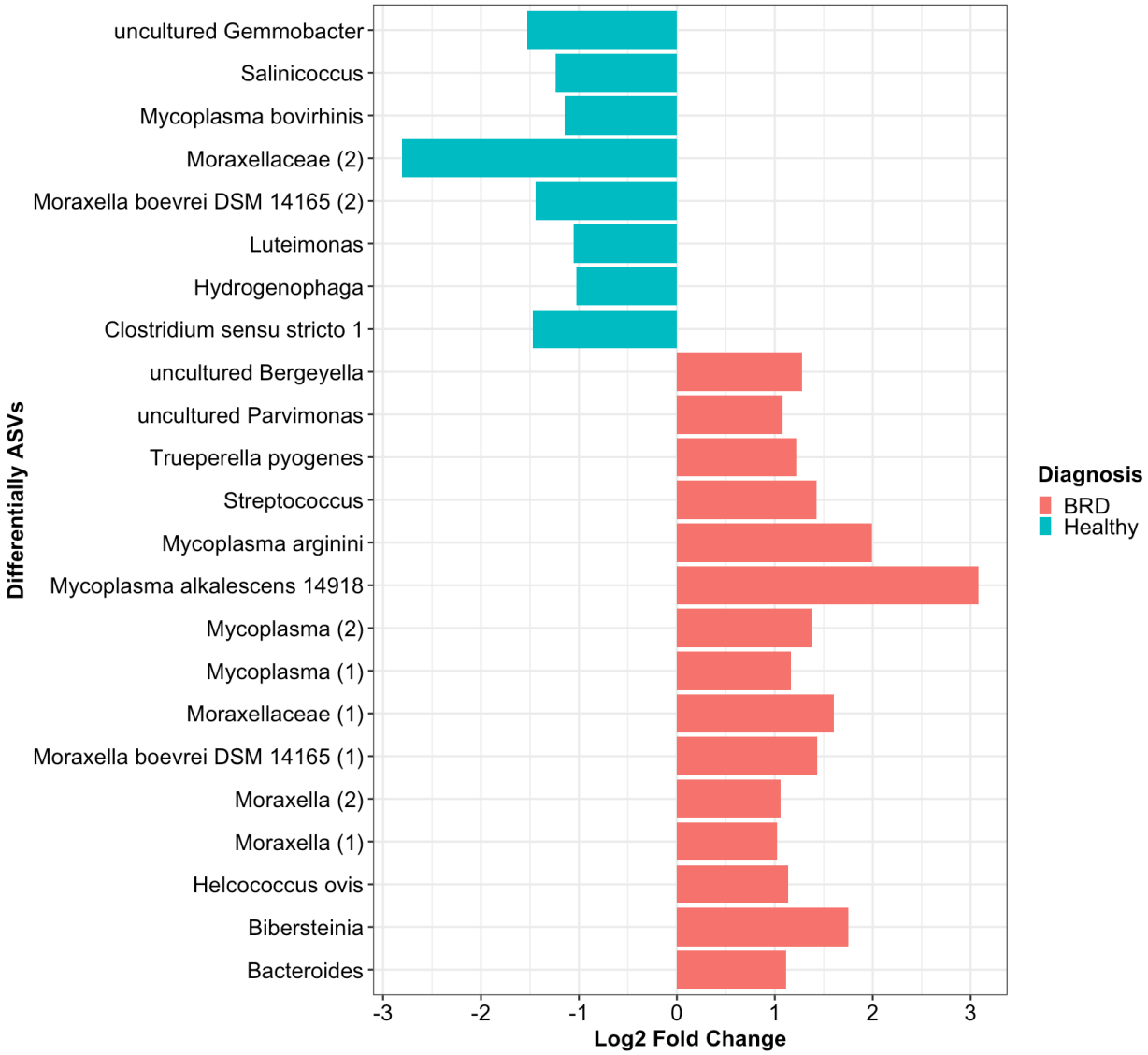

Figure 4.5. Differentially abundant taxa (ASVs) between BRD and healthy animals. Bar plot shows the taxa with a log2 fold change greater than 1 ($P \leq 0.05$). Those with a log2 fold change > 1 were those enriched in BRD animals, while a log2 fold change < 1 were those enriched in the healthy animals. Taxa that contain numbers in parenthesis represent multiple ASVs with the same taxonomy.

4.2 Quantification of BRD-associated species by qPCR

4.2.1 Prevalence of BRD-associated pathogens and relative abundance

Quantification of *Mannheimina haemolytica, Mycoplasma bovis, Pasteurella muloticida, Histophilus somni,* and total 16S rRNA was performed using the DNA extracted from nasal swabs

collected from BRD and healthy pen-mates. The prevalence of the four bacteria in all the samples and between the BRD and healthy animals show that *H. somni* (98%, 129 samples out of 133) and *P. multocida* (97%, 130 out of 133 samples) had higher prevalence regardless of disease diagnosis in all the samples, relative to *M. bovis* (61%, 93 out of 133 samples) and *M. haemolytica* (70%, 81 out of 133 samples) (Fig. 4.6). Also, a difference in prevalence was observed between BRD and healthy animals for *M. haemolytica* (81 and 61%, respectively), and *M. bovis* (74 and 50.7%, respectively) (Fig. 4.6).

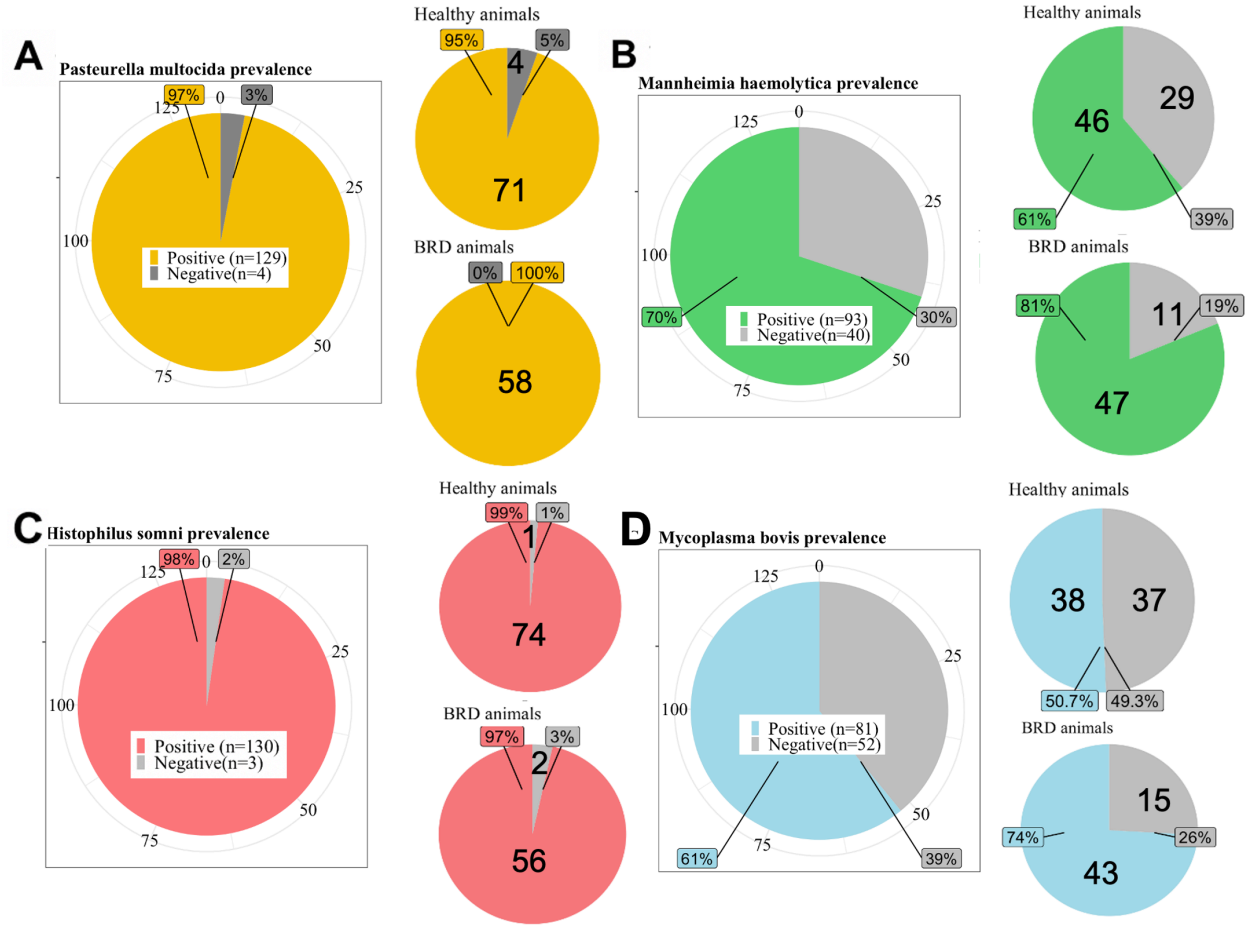

Figure 4.6. Prevalence of *Pasteurella multocida* (A), *Mannheimina haemolytica* (B), *Mycoplasma bovis* (C), and *Histophilus somni* (D) in the nasal microbiota of Holstein steers (N=133) and between healthy (N=75) and BRD (N=58) Holstein steer pen-mates.

4.2.2 16S rRNA gene copy number

Interestingly, the 16S rRNA gene that indicate the absolute bacterial load was significantly higher in the BRD animals than healthy animals (F 1,124.05= 9.5567, $P < 0.002$, Fig. 4.7).

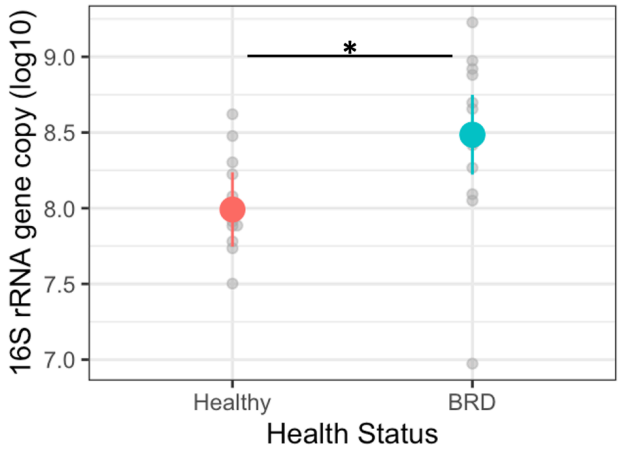

Figure 4.7. Difference in bacterial load (log_{10}) measured by targeting the bacterial 16S rRNA gene between BRD and healthy animals. An asterisk (*) and horizontal line represent a statistical difference ($P \leq 0.05$) between the two groups. Colored circles represent the means of the BRD, and healthy group and the gray dots represent individual samples from each group.

4.2.3 BRD-associated bacteria gene copy number and relative abundance analysis

When analyzing the data from all the animals (N=133), the bacterial load density of *M. bovis* (F 1,122.72= 15.7327, $P < 0.0001$; Fig. 4.8) was increased nearly 100-fold and *M. haemolytica* (F 1,123.25= 10.9789, $P < 0.0001$; Fig. 4.8) was increased nearly 10-fold in BRD compared to healthy animals. On the contrary, *H. somni* bacterial density was about 10% lower in BRD animals (F 1,124.00= 8.9821, $P < 0.003$; Fig. 4.8). No significant difference was observed for *P. multocida* bacterial density based on health status. The abundance of *M. bovis* relative to the total abundance of the four BRD pathogens was significantly higher in BRD than healthy animals (F 1,124.93=16.3239, $P<0.0001$; Fig. 4.8). On the contrary to *H. somni* that presented significant lower relative abundance in BRD than healthy animals (F 1,123.65= 7.4001, $P < 0.007$; Fig. 4.8). No other significance was observed in the bacterial load of *M. bovis, H. somni, P. multocida* and *M. haemolytica* due to date of collection or age.

When analyzing the animals that tested positive for the BRD-related species, load density of *M. bovis* (F 1, 73.951= 11.5278, $P < 0.001$; Fig. 4.9) and *M. haemolytica* (F 1,84.258= 7.8551, $P <$

0.006; Fig. 4.9) were significantly higher in BRD animals than healthy, while *H. somni* was lower in BRD animals (F 1, 121.49=9.2203, $P < 0.002$; Fig. 4.9). No significant difference was observed in *P. multocida* bacterial load and relative abundance values between the two groups. The abundance of *Mycoplasma bovis* relative to the total abundance of the four pathogens significantly increased in the BRD animals than healthy animals (F 1, 75.525= 8.9070, $P < 0.003$; Fig. 4.9). On the contrary, the relative abundance of *P. multocida* significantly decreased in the BRD animals (F 1, 120.23=14.1469, $P < 0.0002$; Fig. 4.9).

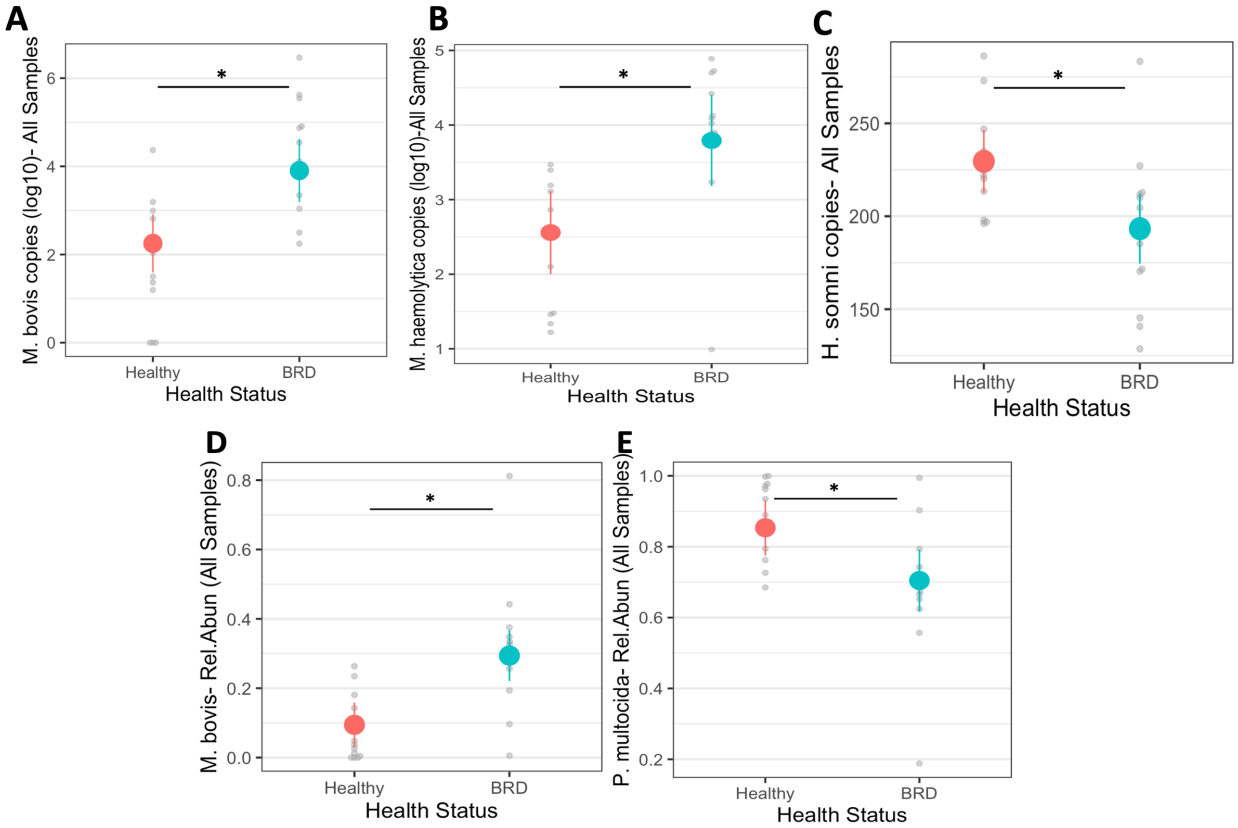

Figure 4.8. Difference in bacterial load for *Mycoplasma bovis, Mannheimia haemolytica* and *Histophilus somni* (A-D) between BRD and healthy animals when including all the samples (n=133). Difference in abundance for *Mycoplasma bovis* and *Pasteurella multocida* relative to the total abundance of the four pathogens (D-E) between BRD and healthy animals when including all the samples. An asterisk (*) and horizontal line represent a statistical difference ($P \leq 0.05$) between the two groups. Colored circles represent the means of the BRD, and healthy group and the gray dots represent individual samples of each group.

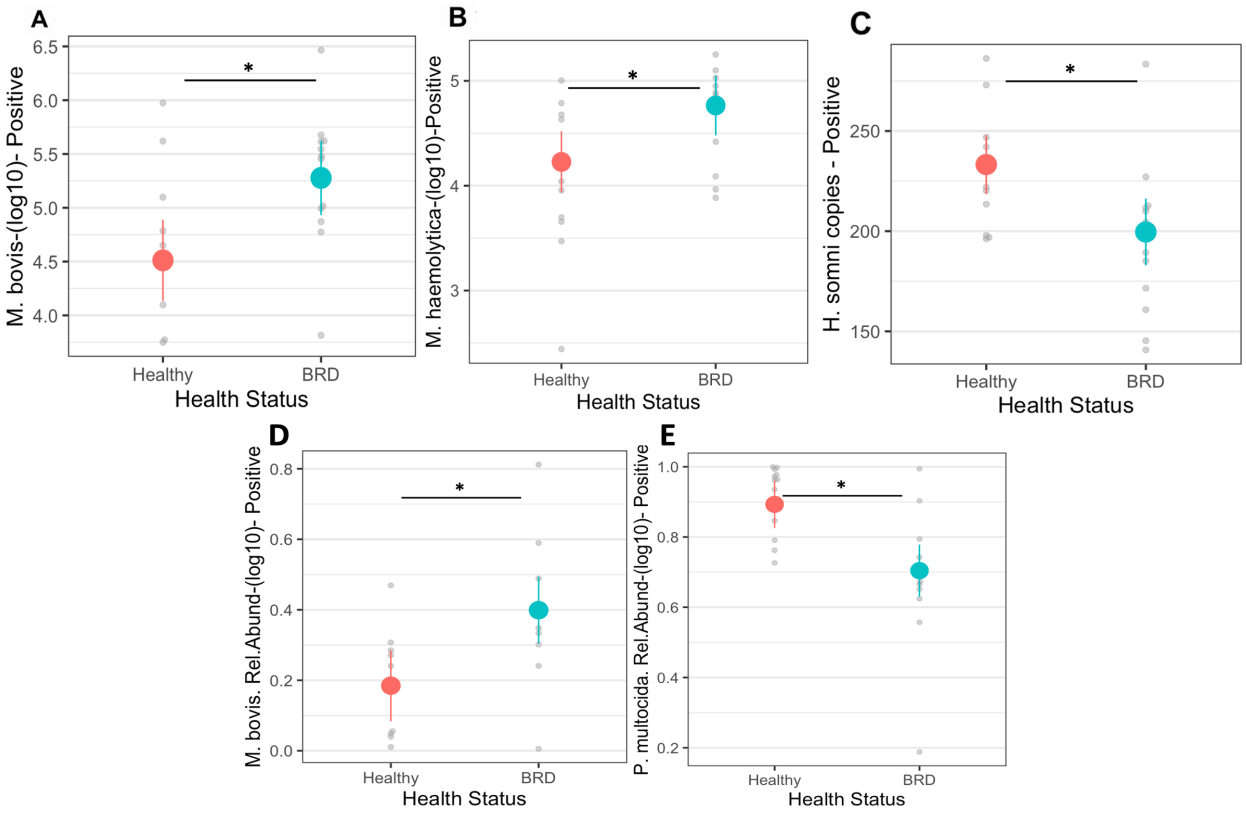

Figure 4.9. Difference in bacterial load for *Mycoplasma bovis, Mannheimia haemolytica* and *Histophilus somni* (A-B) between BRD and healthy animals that tested positive for each bacterium. Difference in abundance for *Mycoplasma bovis* and *Pasteurella multocida* relative to the total abundance of the four pathogens (D-E) between BRD and healthy animals that tested positive for each bacterium. An asterisk (*) and horizontal line represent a statistical difference ($P \leq 0.05$) between the two groups. Colored circles represent the means of the BRD, and healthy group and the gray dots represent individual samples of each group.

Nonetheless, the relative abundance of *M. bovis* and *P. multocida* was significantly affected relative to when the samples were collected in the study (from July to December 2020). In this study, the relative abundance of *M. bovis* was significantly lower in the samples collected around July than the samples collected around December 2020 (F 1, 124.93=16.3239, P <0.00001; Fig. 4.10). On the contrary, *P. multocida* relative abundance was significantly higher in the samples collected close to July than the samples collected close to December (F 1, 128.94= 7.9508, P <0.005; Fig. 4.10). In addition, when analyzing only the samples that tested positive to *M. bovis,* the relative abundance of this bacteria was significantly lower at the beginning of the study than at the end (F 1, 76.899= 8.7658, P < 0.004; Fig. 4.10). As previously reported that the date when the samples were collected was positive associated with the average daily temperature, no

significant differences were observed due to environmental temperature in the relative abundances of *M. bovis* and *P. multocida.*

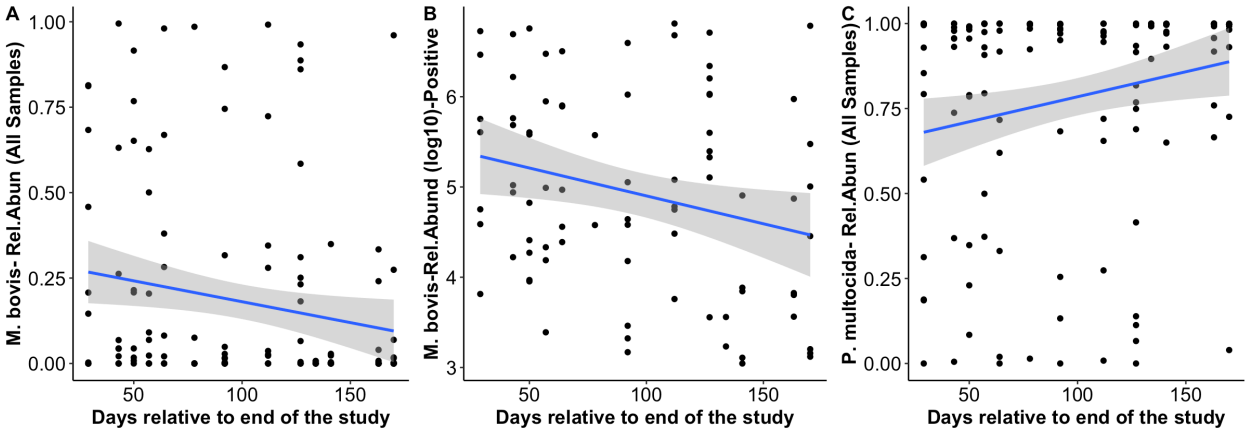

Figure 4.10. Variation in *Mycoplasma bovis* relative abundance in all samples (A), in samples that tested positive (B) and *Pasteurella multocida* (C) relative abundance in all samples (± 95% confident bands) relative to the date of collection. High values in the x axis represent the dates of samples collected at the beginning of the study; low values represent the dates of samples collected at the end of the study.

4.2.4 Linear Discriminant Analysis

In this study, a Linear Discriminant Analysis including BRD-associated pathogens gene copy number, 16S rRNA gene copy number, age, and rectal temperature was used to generate an equation that could be used to discriminate the animals according to diagnosis by the DART method. The BRD-associated bacteria and 16S rRNA gene copy number values were tested first, along with the rectal temperature and age data. However, the number of gene copy numbers was log10 transformed due to a better result in the model's sensitivity and specificity score than using the raw data. Two types of analysis were performed using LDA: 1, including all samples regardless of if the samples tested negative or positive to the four BRD-associated bacteria (n=129) and 2, only including the samples that tested positive for the four BRD associated bacteria (n=66). The best possible model was selected using the following criteria: high sensitivity, specificity, true positive percent (TP%) that indicates the true positive animals identified by LDA model when using the visual observation as the reference, true negatives percent (TN%) that indicates the true

negative animals identified by LDA model when using the visual observation as the reference, and low misclassification rate.

LDA analysis 1: discriminate healthy and BRD animals including all samples (n=129).

Three possible models were selected for this analysis after applying the criteria mentioned above (Equations 3-5); the summary data is presented in Table 3. For model , the LDA identified 23.1% of the animals as TPs (12 out of 19 BRD animals present in the testing set), and 48.1% of the animals as TNs (25 out of 33 healthy animals present in the testing set) and 28.8% of the animals were misclassified when using the visual observations as the reference. For model 2, the LDA identified 23.1% of the TPs animals (12 out of 19 BRD animals present in the testing set), 46.2% of the TNs animals (24 out of 33 healthy animals present in the testing set) and 30.76% of the animals were misclassified. For model 3, the LDA identified 26.9% TPs animals (14 out of 19 BRD animals present in the testing set), 42.3% of the animals as TNs (22 out of 33 healthy animals present in the testing set) and 30.76% of the animals in the testing set were misclassified when using the visual observation as the reference. The equations used to create the models are shown in Table A2. Visual representation of the classification of BRD and healthy animals are shown in Fig. A6.

1. **Model 1**: Rectal temperature (°C) + animal age (month) + *M. haemolytica* $\log_{10}$ gene copy number + *H. somni* $\log_{10}$ gene copy number + *P. multocida* $\log_{10}$ gene copy number + *M. bovis* $\log_{10}$ gene copy number (3)
2. **Model 2:** Rectal temperature (°C) + *M. haemolytica* $\log_{10}$ gene copy number + *H. somni* $\log_{10}$ gene copy number + *P. multocida* $\log_{10}$ gene copy number + *M. bovis* $\log_{10}$ gene copy number + 16S rRNA $\log_{10}$ gene copy number. (4)
3. **Model 3:** *M. haemolytica* $\log_{10}$ gene copy number + *H. somni* $\log_{10}$ gene copy number + *P. multocida* $\log_{10}$ gene copy number + *M. bovis* $\log_{10}$ gene copy number + 16S rRNA $\log_{10}$ gene copy number (5)

Table 3. Linear Discriminant Model summary result of the best possible models, including all samples (n=129), that could be used to discriminate BRD and healthy animals.

Criteria	Model 1	Model 2	Model 3
TP (%)	23.1	23.1	26.9
FN(%)	15.4	17.3	21.2
TN (%)	48.1	46.2	42.3
FP (%)	13.5	13.5	9.6
Misclassification error	0.2884	0.3076	0.3076
Sensitivity (%)	60	57	60
Specificity (%)	78	78	81

TP: True Positives (%)
FN: False Negatives (%)
TN: True Negatives (%)
FP: False Positives (%)

LDA analysis2 : discriminate healthy and BRD animals including only samples that tested positive to BRD-associated pathogen.

In the dataset that include samples positive to the four BRD-pathogens (n=66), after applying the criteria mentioned above, two possible models predicted by LDA were identified (equations 6-7*)* following the criteria mentioned above, summary data is shown in Table 4. Equation used to create the models are shown in Table A3. Model 4, the LDA identified 44.4% of the animals as TPs (12 out of 16 BRD animals present in the testing set). 25.9% of the animals (8 out of 11 healthy animals present in the testing set) and 25.92% of the animals were misclassified when suing the visual observations as the reference. Model 5, the LDA model identified 48.1% of the animals as TPs (13 out of 16 BRD animals present in the testing set), 29.6% of the animals as TNs (8 out of 11 healthy animals present in the testing set) and 22.22% of the animals were misclassified in the testing set when using the visual observation as the reference. Sample classification into BRD and healthy animals are shown in Fig. A7.

1. **Model 4:** Rectal temperature (°C) + *H. somni* $\log_{10}$ gene copy number + *P. multocida* $\log_{10}$ gene copy number + *M. haemolytica* $\log_{10}$ gene copy number + *M. bovis* $\log_{10}$ gene copy number + 16S rRNA $\log_{10}$ gene copy number. (6)
2. **Model 5:** *M. bovis* $\log_{10}$ gene copy number + *P. multocida* $\log_{10}$ gene copy number + 16S rRNA $\log_{10}$ gene copy number. (7)

Table 4. Summary of Linear Discriminant Model result of the best possible models, only including samples that tested positive to the four BRD-associated bacteria that could be used to discriminate BRD and healthy animals.

Criteria	Model 4	Model 5
TP (%)	44.4	48.1
FN (%)	11.1	11.1
TN (%)	25.9	29.6
FP (%)	14.8	11.1
Misclassification error	0.2592	0.2222
Sensitivity (%)	80	81
Specificity (%)	66	72

TP: True Positives (%)
FN: False Negatives (%)
TN: True Negatives (%)
FP: False Positives (%)

From the three models presented including all the samples, the highest sensitivity and specificity between the diagnosis given by the DART method and the qPCR method were 60% and 81%. On the contrary, when only the samples that tested positive for the four bacteria, the specificity and sensitivity of the models increased to 80-81% and 66-72%. One pitfall of these models relies upon the fact that the BRD and healthy animal categorization were identified by detecting observable BRD clinical signs. Thus, there is a chance that BRD classification was inaccurate in this data set, which could lead to a less accurate linear discriminant model. One clear example of the difficulty of diagnosing accurately an animal with BRD was observed in this study. Based on BRD detection using clinical signs, eight animals were classified as healthy. However, based on the LDA model 1, these samples were classified as BRD. Looking closely at the rectal temperature collected from these animals, one animal classified as Healthy, presented a temperature of 103-106°F were most of the healthy animals in the study presented a rectal temperature of 102°F.

Besides, an LDA model was performed to identify which predictors (*Mycoplasma bovis, Mannheimia haemolytica, Histophilus somni, Pasteurella multocida*, 16S rRNA gene copy number, animal rectal temperature, and age) have more impact on the classification of BRD and healthy animals. Interestingly, when rectal temperature, age and *P. multocida* are removed from the model, the misclassification rate increase to 0.307. On the contrary, when 16S rRNA gene, *M, bovis, H. somni* and *M. haemolytica* are removed, the misclassification error decrease to 0.2886 (Table A4). These results suggest that the presence of rectal temperature, age and *P. multocida* in

the model are important to discriminate BRD and healthy animals, whereas the removal of 16S rRNA gene, *M, bovis, H. somni* and *M. haemolytica,* improves the model.

Also, with the removal of temperature but maintaining *M. bovis* in the model, the percent of true positives identified in the dataset increased to 26.9 in comparison to the 17.3 when *M. bovis* is absent but temperature remains in the model (Table A4). These results also suggest that even though the absence of *M. bovis* in the model improves the misclassification error, its presence give more information in discriminating BRD and healthy animals than temperature (26.9 vs 17.3 %).

4.2.5 Nasal microbiome co-occurrence analysis

In the study, a total of 18,010 ASVs were identified in the samples. However, after removing any rare ASVs with an average abundance below 0.0001, 1,236 AVSs remained in the analysis and were used in the co-occurrence analysis. The total number of species pair combinations, species pairs that were not included in the model due to an expected co-occurrence <1, positive, negative, and random pairs association calculated for the BRD and healthy groups is found in Table 5. Positive association indicate the ASVs pairs combination that are more likely to co-exist in the same sample, whereas the negative associations represent the ASVs that are less likely to no co-exist in the same sample.

Table 5. Co-occurrence analysis summary for BRD (n=57) and Healthy (n=74) groups.

Group	ASV pair combinations	ASV pair combinations not included	ASV pairs remained	ASV pair associations (+)	ASV pair associations (-)	Random unclassified ASV pair combinations
BRD	763,230	60,569	702,661	92,500	15,313	594,848
Healthy	763,230	41,783	721,447	147,864	15,177	558,406

After applying the filtering thresholds (ASV pairs present in more than 60% of the samples and a probability greater than 0.9, a total of 280 positive species pairs remained in the healthy group (n=74) and 90 positive species pairs in the BRD group (n=57). In the healthy group, positive ASV pair combinations were mainly composed of bacteria belonging to the phylum *Actinobacteria, Bacteroidetes, Euryarchaeota, Fibrobacteres, Firmicutes, Patescibacteria, Proteobacteria,* and

Spirochaetes (Fig. A8). Only 32 ASV pair combinations present in 73 out of 74 Healthy animals with a probability greater than 0.9, this result demonstrates the ASVs pair combination that are more common to be observed in the nasal microbiome of healthy animals (probability that the two ASVs will be observed in the same sample greater than 0.90, Table 6).

Table 6. Co-occurrence analysis with ASV pair combinations present in 73 out of 74 healthy samples with a probability of occurrence in the same sample greater than 0.9.

ASVs	**Genus_1**	**ASVs**	**Genus_2**
ASV10188	*Romboutsia*	ASV16829	*Guggenheimella*
ASV10188	*Romboutsia*	ASV2838	*Prevotellaceae* NK3B31 group
ASV10188	*Romboutsia*	ASV465	*Prevotellaceae* NK3B31 group
ASV1409	*Georgenia*	ASV13816	*Parapusillimonas*
ASV1409	*Georgenia*	ASV8900	*Uncultured Acidaminococcaceae*
ASV14820	*Acinetobacter*	ASV10188	*Romboutsia*
ASV14820	*Acinetobacter*	ASV16829	*Guggenheimella*
ASV14820	*Acinetobacter*	ASV17312	*Methanobrevibacter*
ASV14820	*Acinetobacter*	ASV2534	*Olsenella*
ASV14820	*Acinetobacter*	ASV2838	*Prevotellaceae* NK3B31 group
ASV14820	*Acinetobacter*	ASV465	*Prevotellaceae* NK3B31 group
ASV16829	*Guggenheimella*	ASV2838	*Prevotellaceae* NK3B31 group
ASV16829	*Guggenheimella*	ASV465	*Prevotellaceae* NK3B31 group
ASV17312	*Methanobrevibacter*	ASV10188	*Romboutsia*
ASV17312	*Methanobrevibacter*	ASV16829	*Guggenheimella*
ASV17312	*Methanobrevibacter*	ASV2534	*Olsenella*
ASV17312	*Methanobrevibacter*	ASV2838	*Prevotellaceae* NK3B31 group
ASV17312	*Methanobrevibacter*	ASV465	*Prevotellaceae* NK3B31 group
ASV2534	*Olsenella*	ASV10188	*Romboutsia*
ASV2534	*Olsenella*	ASV16829	*Guggenheimella*
ASV2534	*Olsenella*	ASV2838	*Prevotellaceae* NK3B31 group
ASV2534	*Olsenella*	ASV465	*Prevotellaceae* NK3B31 group
ASV3244	*Corynebacterium* 1	ASV10188	*Romboutsia*
ASV3244	*Corynebacterium* 1	ASV14820	*Acinetobacter*

Table 6 continued

ASV3244	*Corynebacterium* 1	ASV16829	*Guggenheimella*
ASV3244	*Corynebacterium* 1	ASV17312	*Methanobrevibacter*
ASV3244	*Corynebacterium* 1	ASV2534	*Olsenella*
ASV3244	*Corynebacterium* 1	ASV2838	*Prevotellaceae* NK3B31 group
ASV3244	*Corynebacterium* 1	ASV465	*Prevotellaceae* NK3B31 group
ASV465	*Prevotellaceae* NK3B31 group	ASV2838	*Prevotellaceae* NK3B31 group
ASV676	*Prevotellaceae* NK3B31 group	ASV6770	*Fermentimonas*
ASV8900	Uncultured *Acidaminococcaceae*	ASV13816	*Parapusillimonas*

In the BRD group, the 90 positive ASV pair combinations were mainly composed of bacteria from the phylum *Actinobacteria, Bacteroidetes, Euryarchaeota, Fibrobacteres, Firmicutes, Patescibacteria*, *Proteobacteria* and *Spirochaetes* (Fig. A9). Only eight pairs were identified in 56 out of 57 BRD animals with a probability of 1 (Table 7), these results indicate the ASVs associations that are common to co-occur in the nasal cavity of a BRD-affected animal with a probability greater than 0.9. This is interesting because it demonstrates that there are more ASVs association that are more common to be observed in the Healthy group (32 ASVs associations) than in the BRD group (8 ASVs associations).

Table 7. Co-occurrence analysis with ASV pair combinations present in 56 out of 57 BRD samples with a probability of occurrence in the same sample of 1.

ASV	**Genus_1**	**ASV**	**Genus_2**
ASV11414	*Flavobacterium*	ASV17834	*Halomonas*
ASV5703	*Glutamicibacter*	ASV17834	*Halomonas*
ASV17312	*Methanobrevibacter*	ASV2463	*Gulosibacter*
ASV8765	*Prevotella* 1	ASV465	*Prevotellaceae* NK3B31 group
ASV465	*Prevotellaceae* NK3B31 group	ASV4663	*Prevotellaceae* NK3B31 group
ASV8765	*Prevotella* 1	ASV4663	*Prevotellaceae* NK3B31 group
ASV11	*Clostridium sensu stricto* 1	ASV5609	*Bacteroides*
ASV11414	*Flavobacterium*	ASV5703	*Glutamicibacter*

After applying the rules that negative associations should be present in sites lower than the expected number of sites with a probability <0.01 and with a probability of <0.05 that the two

ASVs will be present in <1 samples, a total of 178 negative ASV pair associations were identified in the healthy group with bacteria members of the phylum *Actinobacteria, BRC1, Bacteroidetes, Chloroflexi, Cyanobacteria, Euryarchaeota, Firmicutes, Patescibacteria, Proteobacteria, Spirochaetes, Tenericutes, Verucomicrobia, Acidobacteria, Synegistetes,* and *Planctomycetes* (Fig. A10). A total of 49 negative associations in the BRD group were identified with bacterial members the *Actinobacteria, BRC1, Bacteroidetes, Chloroflexi, Cyanobacteria, Euryarchaeota, Firmicutes, Patescibacteria, Proteobacteria, Spirochaetes, Tenericutes, Verucomicrobia, Acidobacteria, Fusobacterium* and *Deinococcus-Thermus* phylum (Fig. A11).

In the healthy group, 32 negative ASV pairs associations were identified to less likely to co-exist in the samples with a probability <0.04 (Table 8). In the BRD group, 13 negative ASV pair associations were identified to less likely to co-exist in the samples with a probability <0.04 (Table 9).

Table 8. Co-occurrence analysis showing negative ASV pair combinations that are less likely to co-exist in the heathy group with a probability <0.04.

ASV_1	**Genus_1**	**ASV_1**	**Genus_2**
ASV4603	uncultured *Bacteroides* sp.	ASV11512	uncultured *Prevotellaceae bacterium*
ASV2463	uncultured *Gulosibacter*	ASV9158	*Mycoplasma bovirhinis*
ASV285	uncultured *Bifidobacteriaceae*	ASV11166	*Intrasporangiaceae*
ASV13000	*Prevotellaceae* UCG-003	ASV11166	*Intrasporangiaceae*
ASV14432	uncultured *Guggenheimella*	ASV11166	*Intrasporangiaceae*
ASV14432	uncultured *Guggenheimella*	ASV15485	*Thermomonas*
ASV13000	*Prevotellaceae* UCG-003	ASV15485	*Thermomonas*
ASV285	uncultured *Bifidobacteriaceae*	ASV15485	*Thermomonas*
ASV13000	*Prevotellaceae* UCG-003	ASV1851	uncultured *Desulfocaldus* sp.
ASV285	uncultured *Bifidobacteriaceae*	ASV1851	uncultured *Desulfocaldus* sp.
ASV14432	uncultured *Guggenheimella*	ASV1851	uncultured *Desulfocaldus* sp.
ASV1706	*Corynebacterium* 1	ASV11166	*Intrasporangiaceae*
ASV14362	uncultured *Lysinibacillus*	ASV11166	*Intrasporangiaceae*
ASV1706	*Corynebacterium* 1	ASV15485	*Thermomonas*

Table 8 continued

ASV14362	uncultured *Lysinibacillus*	ASV15485	*Thermomonas*
ASV14362	uncultured *Lysinibacillus*	ASV1851	uncultured *Desulfocaldus* sp.
ASV1706	*Corynebacterium* 1	ASV1851	uncultured *Desulfocaldus* sp.
ASV7028	*Arthrobacter*	ASV11166	*Intrasporangiaceae*
ASV247	uncultured *Oligella*	ASV11166	*Intrasporangiaceae*
ASV12358	uncultured *Cellvibrio*	ASV1237	*Microbacteriaceae*
ASV7028	*Arthrobacter*	ASV15485	*Thermomonas*
ASV247	uncultured *Oligella*	ASV15485	*Thermomonas*
ASV247	uncultured *Oligella*	ASV1851	uncultured *Desulfocaldus* sp.
ASV7028	*Arthrobacter*	ASV1851	uncultured *Desulfocaldus* sp.
ASV10880	*Methanobrevibacter*	ASV11196	uncultured *Chthoniobacteraceae*
ASV10880	*Methanobrevibacter*	ASV13602	*Paracoccus*
ASV7532	*Mycoplasma bovirhinis*	ASV15514	*uncultured Synergistaceae*
ASV10880	*Methanobrevibacter*	ASV16570	*Flavobacterium*
ASV5038	*Aerococcus*	ASV4856	*Prevotellaceae*
ASV7532	*Mycoplasma bovirhinis*	ASV5038	*Aerococcus*
ASV7532	*Mycoplasma bovirhinis*	ASV6977	*Taibaiella*
ASV804	*Ruminococcus* 1	ASV9546	*Vitreoscilla*
ASV10259	*Chishuiella* sp. YIM 102668	ASV11196	uncultured *Chthoniobacter*
ASV10259	*Chishuiella* sp. YIM 102668	ASV16570	*Flavobacterium*
ASV10259	*Chishuiella* sp. YIM 102668	ASV4856	*Prevotellaceae*
ASV7754	*Paludibacter*	ASV13602	*Paracoccus*

Table 9. Co-occurrence analysis showing negative ASV pair combinations that are less likely to co-exist in the BRD group with a probability <0.04.

ASV_1	**Genus_1**	**ASV_1**	**Genus_2**
ASV10164	uncultured *Absconditabacteriales* (SR1)	ASV11196	uncultured *Chthoniobacter*
ASV10164	uncultured *Absconditabacteriales* (SR1)	ASV15485	*Thermomonas*
ASV10164	uncultured *Absconditabacteriales* (SR1)	ASV16570	*Flavobacterium*
ASV15050	uncultured *Jeotgalibaca*	ASV222	*Mycoplasma*

Table 9 continued

ASV7680	*Corynebacterium* sp. C3	ASV7532	*Mycoplasma bovirhinis*
ASV1344	*Planococcaceae*	ASV7532	*Mycoplasma bovirhinis*
ASV13607	*Sphingobacterium jejuense*	ASV7532	*Mycoplasma bovirhinis*
ASV7716	*Weissella paramesenteroides*	ASV1143	uncultured *Murdochiella*
ASV3751	*Mycoplasma bovoculi* M165/69	ASV2127	*Moraxella boevrei* DSM 14165
ASV4603	uncultured *Bacteroides* sp.	ASV222	*Mycoplasma*
ASV12541	uncultured *Rikenellaceae* RC9	ASV3453	*Deinococcus*
ASV10188	*Romboutsia*	ASV7532	*Mycoplasma bovirhinis*
ASV12541	uncultured *Rikenellaceae* RC9	ASV8878	*Moraxellaceae*

In addition to identifying the most prevalent ASV pair combinations in the nasal microbiome of healthy and BRD animals, ASV pair combination between commensal microbiota and BRD-pathogens were identified with more relaxed filtering criteria. This allowed us to detect how common the BRD-pathogens are associated with other bacteria in the cattle nasal cavity. In the healthy group, 96 positive ASV pair combinations; however, the number of samples where these combinations were observed ranged between 5-30 samples out of 74, with a probability between 0.033-0.34 (Table A5). Also, 74 negative ASV pair combinations were observed between *Mycoplasma bovis* and bacteria present in the nasal cavity. These associations were observed in a total of 1-26 samples out of 74 with a probability between 0.044-0.383 (Table A6). Three negative ASV pair combinations were observed in the Healthy group between *Pasteurella multocida* and *Christensenellaceae* R-7 group, *Lachnospiraceae*, and one uncultured *Moheibacter.* They were present in approximately 10 to 14 samples with a probability ranging between 0.15 and 0.21. No other positive or negative bacterial association with *Pasteurella multocida, Histophilus somni*, and *Mannheimina haemolytica* was observed in the healthy animals.

In the BRD group, a total of 43 positive ASV pair combinations between *Mycoplasma bovis* and bacteria in the nasal cavity were observed. These associations were observed in a total of 4-35

samples out of 57 total samples with a probability between 0.025 and 0.547 (Table A7). Also, only one pair combination was observed between *Pasteurella multocida* and *Escherichia-Shigella* in 55 samples with a chance of occurring in the same sample of 0.948. Also, a total of 96 negative ASV pair combinations were observed in the BRD group with presence in approximately 1-27 samples out of 57 total samples with a probability between 0.04 to 0.524 (Table A8). One negative association was also observed between *Pasteurella multocida* and an uncultured *Tissierella sp.* In 10 samples with a probability of 0.203. No other bacterial association with *Histophilus somni* and *Mannheimina haemolytica* was observed in the BRD animals.

CHAPTER 5. DISCUSSION AND CONCLUSIONS

5.1 Discussion

Bovine Respiratory Disease (BRD) is ongoing health and economic problem to the dairy and beef industry and it is mostly developed by multiple factors that make animals susceptible to BRD (Chirase & Greene, 2000; Snowder et al., 2006). Four major pathogens had been identified to have a relation with BRD development: *Pasteurella multocida, Histophilus somni, Mannheimia haemolytica,* and *Mycoplasma bovis* (Klima et al., 2014; Mosier, 2014). Thus, in this study, nasal samples collected from healthy and BRD-affected animals diagnosed by the detection of clinical signs were used to identify a difference in the nasal microbiome between the two groups, prevalence and bacterial load density of the four BRD-associated pathogens, and any interaction of co-occurrence between the commensal nasal microbiome or between the BRD-associated pathogens and nasal microbiome.

Nasal microbiome between BRD and healthy animals

The 16S rRNA gene sequencing is one of the most common methods used to study the bacterial community and composition within a set of samples. The 16S rRNA gene is common and conserved among bacteria; thus, it can be used as a target to identify culturable and unculturable bacteria in a different set of samples. Hence, this approach allowed us to visualize the full picture of the nasal microbiome community present within a sample or between two communities, identify different community traits (e.g., microbial richness, evenness) that could be used to discriminate BRD and healthy animals, and identify differentially abundant taxa between BRD and healthy animals The 16S rRNA gene is common and conserved among bacteria; thus, it can be used as a target to identify culturable and unculturable bacteria in a different set of samples (Gupta et al., 2019; Janda & Abbott, 2007; Sontakke et al., 2009). Also, it's possible to calculate the alpha diversity that measures community richness; evenness, and the phylogenetic relationship found within the samples (Gupta et al., 2019; Sontakke et al., 2009).

Interestingly, in this study, BRD- animals presented lower alpha diversity than healthy animals. Other studies found a similar decrease in alpha diversity, in BRD animals compared to healthy

animals (Timsit et al., 2018; Holman et al., 2015a) In our study, we identified a decrease of approximate 20% in the richness of BRD animals compared to healthy animals. A similar trend was observed in Timsit et al (2018) in which the richness median of BRD animals decrease almost 50% than the median of the healthy animals and in Holman et al. (2015a), the richness of the BRD animals decrease approximate 52% when compared to the healthy animals. Also, is important to mention that these two studies collected nasopharyngeal samples and no nasal swabs, but it interesting that the same pattern (decrease in richness in BRD-affected animals) is observed regardless of the type of sample collected. In our study, it was identified that BRD animals presented a decrease of approximately 11% in the phylogenetic diversity when compared to healthy animals. There is evidence that a greater phylogenetic diversity confers more stability to the ecosystem and resistance to pathogen colonization (Cardinale et al., 2012; Knapp et al., 2017). Ecosystem stability is determined by the resistance, resilience, and functional redundancy that the microbial community could possess (Bissett et al., 2013). Resistance refers to the capacity of a system to change following a disturbance and resilience refers to the capacity of a system to return to its original state after perturbation (Pimm, 1984). Functional redundancy refers to the number of taxa present in a community performing a given or similar function (Konopka et al., 2015). Additionally, a mechanism in which diversity confers stability to the ecosystem includes statistical averaging and compensatory dynamics (Cardinale et al., 2012). Statistical averaging refers to a variation in the abundance of different species that decreases the variability of aggregating more ecosystem variables (Doak et al., 1998). Compensatory dynamics refers to the competitive interactions or differential responses of an ecosystem to environmental fluctuations that could lead to asynchrony in their responses (Tilman et al., 2001). Thus, it could be hypothesized that the greater phylogenetic diversity, richness, and evenness observed in healthy animals provided greater microbial stability to the environment which conferred the capacity to recover from disruption and compensate the loss of a member that perform a common function (functional redundancy), a capacity that could have decreased in the microbial community of BRD animals. Additionally, a combined effect of statistical averaging (i.e., increase in the abundance of particular community members while avoiding the aggregation of other members) conferred a different compensatory dynamic response (e.g. how the system respond to distributions) could have resulted in susceptibility to pathogen colonization in the BRD animals than healthy animals; nevertheless, more research is needed to test this hypothesis.

Additionally, differences in the alpha diversity were observed in this study relative to when the samples were collected with higher alpha diversity in samples collected in July 2020 and it decreases towards the end of the study (December 2020). Because the region in which the samples were collected changes in environmental temperature and humidity throughout the year, higher temperatures in summer (July) than winter (December), the change in the alpha diversity could be a result of the animal susceptibility or stress in response to changes in environmental temperatures. Several studies had identified that temperature has an impact on the gut bacterial community leading to a decrease in diversity or specific changes in bacterial members such as Firmicutes (Hylander & Repasky, 2019; Sepulveda & Moeller, 2020). Thus, it is possible that the temperature could contribute to the low alpha diversity values in the samples collected towards the end of the study.

Taxonomical composition of the nasal samples collected from BRD and healthy animals were mostly composed at phylum level of Proteobacteria, Firmicutes, Bacteriodetes, and Actinobacteria regardless of the health status and agrees with previous studies that had been identified as common nasal microbiota members regardless of the health status (McMullen, et al., 2020). However, differences at the Genera level were observed between BRD and healthy animals. Healthy animals presented bacteria genera *Clostridium sensu stricto* 1*,* members of the *Moraxellaceae* family, the genera *Mycoplasma bovirhinis* and *Moraxella boevrei* DSM 14165. In a previous study analyzing the microbial community of different sites in the upper and lower respiratory of healthy cattle, the presence of *Moraxella, Mannheimia, Clostridium senso stricto* 1*,* and *Mycoplasma* were identified as members of the cattle's nostrils microbiome (McMullen et al., 2020). Our results show that *Mycoplasma* sp. was significantly enriched in BRD animals, results that are in agreement with findings in previous studies (McDaneld et al., 2018). This study collected nasal swabs from animals diagnosed with BRD and from healthy cohorts. The group identified that in the BRD-affected animals, the genera *Mycoplasma* sp, *Psychrobacter* sp and *Mannheimia* sp. were the most predominant in comparison with the healthy group. Also, in this study, *Truperella pyogenes,* the genera *Bibersteinia, Streptococcus* and *Moraxella* significantly increased (1.230, 1.754, 1.42 and 1 log2 fold change) in the BRD compared to healthy animals. Other studies identified *Truperella pyrogenes* as a secondary pathogen for BRD, and *Bibersteinia* strains that can produce leukotoxin tend to be related to pneumonia cases in bighorn sheep (Dassanayake et

al., 2010; Klima et al., 2014; Pardon & Buczinski, 2020). Also, the species *Bibersteinia trehalosi* was isolated from the lung of cattle that died of acute pneumonia (Cortese et al., 2012), suggesting that this species could also be related to BRD cases. Also, the genera *Moraxella* and *Streptococcus* had been identified as a commensal member of the cattle's nostril and nasopharynx (McMullen, et al., 2020). A *Mycoplasma* species (*Mycoplasma bovirhinis)* and *Clostridium sensu stricto* were also enriched in the healthy animals; nonetheless, these bacteria have been categorized as commensal or opportunistic bacterial pathogens (Angen et al., 2009; Pardon & Buczinski, 2020, McMullen et al., 2020). These results provide more insight into possible bacteria that could be related to the development of BRD and that could be used in the process of BRD diagnosis.

Prevalence and Quantification of BRD-associated bacteria and bacterial 16S rRNA gene

Characterization of the nasal microbiome community provides knowledge of the community structure present in animals diagnosed with BRD compared to healthy animals. Nonetheless, this method has low phylogenetic power to identify the specific species taxonomical levels and poor discrimination of some particular genera (Janda & Abbott, 2007). A different approach to identify specific species within the samples is the real-time quantitative polymerase chain reaction (qPCR). This method's advantage is that bacterial species can be targeted and quantified in different types of samples (Fulton & Confer, 2012) and the sensitivity of the method is much higher than 16S rRNA gene amplicon sequencing. This method was used to quantify BRD-associated species *Mannheimina haemolytica, Mycoplasma bovis, Pasteurella multoicida, Histophilus somni*, as well as the quantification of the total bacterial community determined by bacterial 16S rRNA gene between BRD and healthy animals.

In this study, the16S rRNA gene abundance was significantly increased in the BRD animals than the healthy animals. This could be an indication that the bacterial load in nasal cavity of BRD is higher than the healthy animals or that a larger mucosal sample was collected from sick animals. Also, *Pasteurella multocida, Histophilus somni* were present in nearly all the samples regardless of health status, whereas species *Mannheimina haemolytica* and *Mycoplasma bovis* were more prevalent and presented higher load density in the BRD animals than healthy animals. A similar result was observed in a study in which nasal swabs were collected only from healthy animals, and qPCR was used to target *H. somni, P. multocida,* and *M. haemolytica* (Thomas et al., 2019). In

that study, *H. somni, P. multocida* were the bacteria with the highest prevalence in the samples (75.6% and 26.8% among all samples); whereas *M. haemolytica* was only present in 5.7% of the samples. These results suggest that *H. somni and P. multocida* could be part of the core nasal microbiome (McMullen et al., 2020). Nonetheless, these bacteria are still considered BRD pathogens. In the case of *P. multocida,* this bacterium has been linked with respiratory disease in dairy calves due to isolation from lung tissues, nasopharyngeal swabs, and transtracheal washes from sick animals and there is strong evidence of its relation with cattle's shipping fever (Dabo et al., 2008). In a study, *P. multocida* was the bacterium most isolated from trans-tracheal aspiration samples of BRD-affected cattle (54.8%) followed by *M. haemolytica* (30.5%) and lastly *H. somni* (22.9%) (Timsit et al., 2017). *H. somni* has been isolated from 10% of the lungs of animals that died of BRD and it has been also associated to produce infection in conjunction with the bovine respiratory syncytial virus (BRSV) (Fulton 2003; Agnes et al., 2013). In future studies, it will be important to determine the causative agent of BRD to determine if the nasal microbiome composition indicates the causative agent in the lung.

Another finding in this study is that *M. haemolytica* and *M. bovis* presented higher prevalence and load density in BRD animals than healthy animals and could be associated with disease development. In previous studies, *Mycoplasma bovis* and *Mannheimia haemolytica* have been identified to play a crucial role in chronic pneumonia in cattle. Both of these bacteria are recognized as commensal members of the nasal microbiome, and in situations where the host defense is compromised by stress, bacterial or viral infection, they can access and colonize the lung (Caswell & Archambault, 1996; Rice et al., 2008). Interestingly, *M. bovis* has been identified as a primary pathogen for BRD (Pardon & Buczinski, 2020). Timsit et al (2018) identified that *M. bovis* was significantly enriched in deep nasopharyngeal and transtracheal aspiration from BRD-affected cattle. Also, *M. haemolytica*, had been identified in nasopharyngeal swabs and lung tissue samples of animal that died of acute fibrinous pneumonia (Klima et al., 2014). The authors identified the presence of *Mannheimia* sp. in 91% of the cattle sampled in the study, followed by *Mycoplasma bovis* (63%) and *H. somni* (57%). An interesting finding in this study is that combinations between *Mycoplasma bovis, Mannheimia* sp. and Bovine Viral Diarrhea Virus (BVDV) were found in 97% of the BRD cases. With these results, it is evident that *M. bovis* and *M. haemolytica* play a role in BRD development; unfortunately, the majority of these studies did not

collect nasal samples to identify the presence of the BRD-associated bacteria. In our study, since both *M. bovis* and *M. haemolytica* were enriched in the sick animals, it is possible that they both infected the lung, but this hypothesis would need to be tested in future animals to draw a connection between the lung and nasal microbiomes.

Linear Discriminant Analysis (LDA) to discriminate BRD and healthy animals

Bovine Respiratory Disease (BRD) development is associated with multiple factors (i.e., predisposing, environmental, and epidemiological factors), making its diagnosis and treatment a challenge for producers. Visual observation of animals with BRD symptoms is the method that producers use to diagnose an animal with BRD. However, not all animals present the symptoms with the same intensity. Evidence demonstrates the diagnosis based on visualization of BRD clinical signs possess low sensitivity (62%) in correctly identifying individuals with the disease and low specificity (63%) in identifying animals without the disease (Loong, 2003; White & Renter, 2009; Wolfger et al., 2015). For this reason, a Linear Discriminant Analysis including the BRD-associated bacteria load density, 16S rRNA gene density, rectal temperature, and animal age were performed to discriminate between BRD and healthy animals. Five models were identified in this study with a sensitivity ranging between 60-80% and specificity ranging between 66-82%. However, one pitfall of this approach is that the animals used to create the model were diagnosed according to their visual symptoms; thus, there is a possibility of having misdiagnosed animals in the study.

Nasal Microbiome Co-occurrence

Bacteria present in an environment will interact with one another, and depending on the signal detected, can change their behavior on a population-wide scale (Waters & Bassler, 2005). Because of this behavior, there is a hypothesis that commensal members of the nasal microbiome could play a role in the host health by decreasing or enhancing the chance of pathogenic bacterial colonization (Corbeil et al., 1985; Timsit et al., 2018). Most BRD studies had focused only on identifying the microbial community as predicted by 16S rRNA gene sequencing and quantification of BRD-associated pathogens present in the nasal cavity of BRD and healthy animals. Few studies had looked into nasal microbiome bacterial interactions or associations. Thus,

a co-occurrence analysis was performed to identify any ASV pairs specific to the BRD and healthy groups.

In this study, it was identified that the nasal microbiome of healthy animals presented more positive associations (species pairs that co-exist in the same sample) and negative associations (species pairs that do not co-exist in the same sample) than the BRD animals. It has been shown that microbes that often coexist close to each other, increase the likelihood to interact with each other (Gentry et al., 2015). Thus, the difference in the number of co-occurrence could be attributed to the higher diversity observed in the healthy animals than BRD in this study. Nonetheless, more research is needed to determine the type of co-occurrence or interaction present in the nasal cavity of BRD and healthy animals (e.g., commensalism, synergism, competition, amensalism, and predation) (Gentry et al., 2015) that could provide more information regarding how BRD develops. Also, as previously described, a more diverse community provides a more stable environment to the community (Cardinale et al., 2012; Knapp et al., 2017). Thus, the higher diversity in the healthy than BRD animals provides the community the capacity to recover from disruption and while maintaining the functional redundancy after an alteration, characteristics that could have decreased in the BRD nasal microbiome. On the contrary, the fewer associations observed in the BRD groups could be associated with microbial dysbiosis, indicating an alteration in the microbiota composition as a result of perturbations of the environment (Zeineldin et al., 2020b). Even though microbial dysbiosis in the nasal cavity is associated with BRD, it could not be specified as the causality of the disease (Faner et al., 2017).

From this study, the nasal positive and negative co-occurrence in the healthy and BRD animals was mostly between bacteria from the phylum *Actinobacteria, BRC1, Bacteroidetes, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Euryarchaeota, Fibrobacteres, Firmicutes, Patescibacteria, Proteobacteria Spirochaetes, Tenericutes, Verrucomicrobia, Acidobacteria, Synegistetes,* and *Planctomycetes*. As previously described, the phylum *Proteobacteria, Firmicutes, Bacteroidetes, Tenericutes, Actinobacteria, Fusobacteria, Verrucomicrobia, Cyanobacteria, Chlorflexi*, and *Spirochaetae* compose the upper respiratory tract of cattle regardless of the health status (Zeineldin et al., 2020b). Only one study had look into microbial co-occurrence patterns present in BRD and healthy animals. In this study, the bacterial genera

Mannheimia and *Histophilus* were substantially linked to BRD animals and the genera *Fusobacterium* and *Caviibacter* were associated with BRD-affected animals (Zeineldin et al., 2020b). In our study, we did not find common associations of the nasal microbiome with BRD-associated pathogens. Also, Amat et al., (2019a) collected Nasopharyngeal samples from cattle transported to the feedlot. In this study, negative associations between *Moraxella, Corynebacterium, Methanobrevibacter*, and *Pasteurella* with *Mycoplasma* were observed. Nonetheless, in our study, we found positive associations between *Mycoplasma bovis* and *Corynebacterium* (samples with observed co-occurrence: 17 and 23, and probability of 0.153 and 0.236). No other associations with previously identify species pairs associations (Amat et al., 2019a) were observed in the study. Another interesting result in this study is that bacteria genera *Acinetobacter, Methanobrevibacter,* and *Corynebacterium* 1 that were associated with other members of the nasal microbiome in healthy, had been identified as the cattle nostril members (Amat, et al, 2019a; Holman et al., 2015b; McMullen et al., 2020). In the case of association in the BRD animals, bacteria genera *Flavobacterium, Methanobrevibacter, Clostridium sensu stricto 1,* and *Bacteroides* were the bacteria most likely to be associated with nasal microbiome members and again, these bacteria had been identified a member of the core nasal microbiome in cattle (McMullen et al., 2020).

Additionally, what was surprising in this study, is that no BRD-associated bacteria were identified to have common associations with members of the nasal microbiome. From the four BRD-associated bacteria, *Mycoplasma bovis* was the bacterium that presented associations in the nasal cavity of healthy and BRD animals; however, positive associations were only observed between 5-30 samples out of 74 in the healthy animals and 4-35 samples out of 57 in the BRD group. Only one positive and negative association was observed between *Pasteurella multocida* with *Escherichia-Shigella* and *Tissierella* sp. Three possible ideas could explain the absence of co-occurrence between BRD-associated bacteria and members of the nasal microbiome. The first idea focuses on the fact that sometimes one bacterial member can enhance or avoid the growth of another bacteria (Gentry et al., 2015). Thus, it is possible that members of the nasal microbiome community are not enhancing or interacting with BRD-associated members. In a study, performing cultures of BRD-associated bacteria with other bacteria, it was observed that the presence of specific species like *S. epidermis, Rhodococcus* sp, *Moraxella* sp, *Coryne* sp, *C,*

murium, *Micrococcus*, and *S. viridans* enhanced the growth of *P. multocida, H. somni,* and *M. haemolytica* (Corbeil et al., 1985); nonetheless, no other studies had looked into bacterial interaction and growth in the nasal cavity of healthy and BRD animals. A second idea proposed is that BRD-associated bacteria present a neutralism interaction, which means that two spacially separated population that have no interaction (Gentry et al., 2015). The last idea focused on the fact that BRD-associated bacteria could produce metabolites that decrease their capacity to co-exist with other commensal members. As an example, *Mycoplasma bovis* can produce multiple immunomodulatory responses against macrophages, neutrophils, lymphocytes, and cytokine secretion that could affect the host cells (Rosengarten et al., 2000). *Mannheimia haemolytica* present virulence factors such as adhesin, capsular polysaccharide, fimbriae, iron-regulated outer member proteins, leukotoxin, lipopolysaccharide (LPS), lipoproteins, neuraminidase, sialoglycoprotease, and transferrin-binding proteins that help to overcome host immune response (Rice et al., 2008). *Pasteurella multocida* present other types of virulence factors such as adherence, colonization factors, iron-regulated and acquisition proteins had been identified (Dabo et al., 2008). Nonetheless, these are some examples of virulence factors that affect the host immune response; unfortunately, no studies had looked into how the presence of these pathogens and generation of metabolites could affect the growth or presence of the nasal microbiome in cattle.

Contribution of this work to the field of Bovine Respiratory Disease research

In the field of bovine respiratory disease research, most of the studies working with the nasal microbiome had focused on characterizing the respiratory microbiome using Next-generation sequencing of 16S rRNA gene PCR amplicons (Holman et al., 2015; McMullen et al., 2019; Timsit et al., 2018) or the quantification of BRD-associated pathogens in the respiratory tract (Thomas et al., 2019; Klima et al., 2014). However, few or almost no studies had combined the benefits of the two approaches in characterizing the nasal microbiome and quantifying the presence of the BRD-associated bacteria and few of them had looked into the interactions that might be occurring in the nasal microbiome. One positive outcome of the study presented is that it combines the two approaches which give insight to the full picture of the nasal microbiome composition in BRD-affected cattle and healthy pen-mates and the bacterial load density of the BRD-associated bacteria present in the samples. Even though the BRD-associated bacteria were not differentially abundant

in the nasal cavity of BRD-affected cattle present in this study, by the use of specific qPCR, it was possible to identify that *M. haemolytica* and *M. bovis* were the bacteria most prevalent and abundant in the nasal cavity whereas *P. multocida* and *H. somni* were present in all the samples regardless of the health status. In conclusion, it was possible to detect the presence of BRD-pathogens by collecting nasal swabs from BRD-affected and healthy animals, a type of sample that could be used in the process of developing a method to diagnose animals with BRD.

5.2 Conclusions

In this work, bacterial DNA was extracted from nasal swabs collected from animals diagnosed with BRD and healthy pen-mates. The objectives of this study focused on, i) identify the nasal microbial community diversity between BRD and healthy animals and taxonomical composition through 16S rRNA gene sequencing, ii) detect and quantify the BRD-associated species *Mannheimina haemolytica, Mycoplasma bovis, Pasteurella multoicida, Histophilus somni*, and quantification of the absolute bacterial community determined by bacterial 16S rRNA gene and use the bacterial load values to generate a linear discriminant model that could be used to classify BRD and healthy animals, iii) perform a co-occurrence analysis to identify bacterial ASV pairs that are more prevalent in BRD or healthy animals.

Bacterial 16S rRNA gene sequencing demonstrated that BRD animals present lower alpha diversity values (richness and evenness) than healthy animals. Also, healthy animals showed higher phylogenetic diversity than BRD, which could be translated to healthy animals having a more stable community that resists colonization of pathogen than the microbial community present in BRD animals. Nasal microbiota composition was mostly composed of bacteria from the Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria phylum regardless of the health status. Nonetheless, specific changes at genera and species level, *Mycoplasma, Truperella pyrogenes, Bibersteinia* in BRD, and *Mycoplasma bovirhinis* and *Clostridium sensu stricto* in healthy animals marked a difference between BRD and healthy animals. Unfortunately, no BRD-associated pathogens were identified as differentially abundant in BRD animals than healthy animals by 16S rRNA amplicon sequencing.

Detection and quantification of BRD-associated pathogens revealed that *Histophilus somni* and *Pasteurella multocida* are primarily present in the nasal microbiome regardless of health status. However, the prevalence of *Mannheimina haemolytica* and *Mycoplasma bovis* was more highly prevalent in BRD animals, demonstrating a possible connection between the presence of these two bacteria and BRD development. Also, BRD animals presented higher total bacterial load than healthy animals. The linear discriminant analysis, including BRD-associated pathogens gene copy number, animal age, and rectal temperature proposed in this study, classified BRD and healthy animals but with few improvements in the sensitivity and specificity compared to detection of BRD based on clinical signs visualization. Nevertheless, this method could improve its detection if the model is constructed using true BRD and true healthy animals.

Co-occurrence analysis demonstrated that the bacteria considered core members of the cattle nasal microbiome are the most common to interact and co-occurred regardless of the health status. On the other hand, BRD-associated bacteria are less likely to have associations with the members of the cattle nasal microbiome.

The application of this study can be extended further by the following three steps i) expanding the list of other microbes like viruses, bacteria, fungi that could be related to BRD development (e.g., *Trueperella pyrogenes* and *Bibersteinia)* ii) generation of a model (LDA) that could be used in addition to BRD diagnosis using clinical signs that could increase the accuracy in correctly classifying BRD and healthy animals, iii) detection of possible co-occurrence between BRD-associated pathogens and members of the nasal microbiome.

APPENDIX A

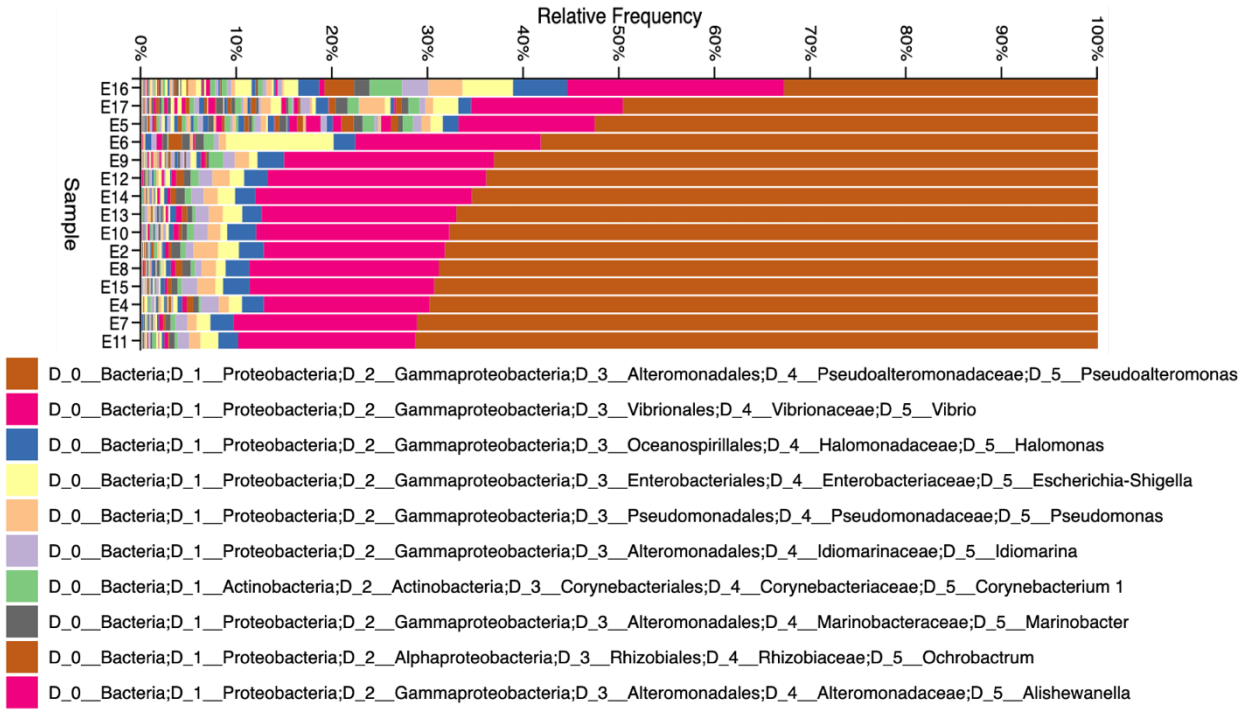

Figure A1. Taxonomical composition showing the top 10 Genera present in the empty tubes used as negative control in the DNA extraction and sequencing step.

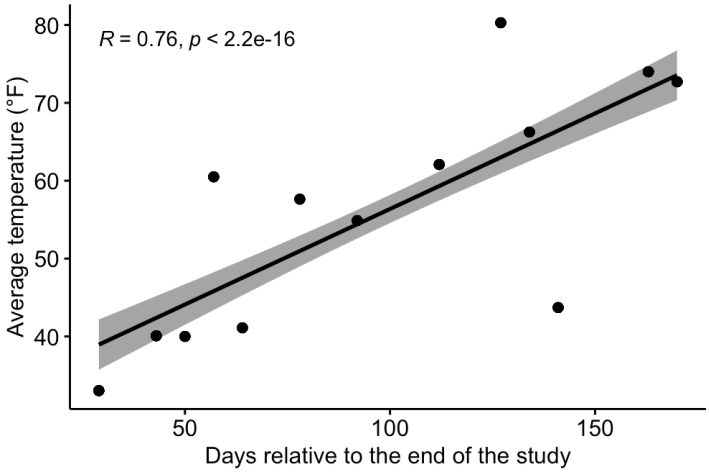

Figure A2. Pearson's correlation with the days relative to the end of the study and average temperature (°F) of when the samples were collected. High values in the x axis represent the dates of samples collected at the beginning of the study; low values represent the dates of samples collected at the end of the study.

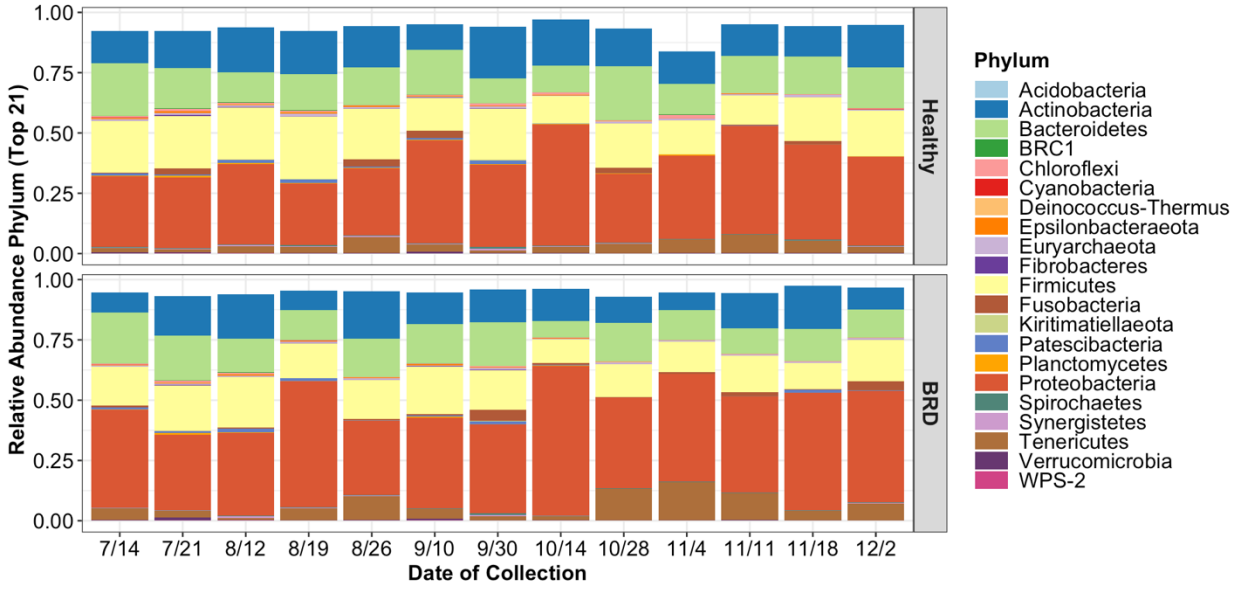

Figure A3. Stacked bar graph showing the average relative abundance of the top 21 phylum present in the nasal microbiome of BRD and healthy animals.

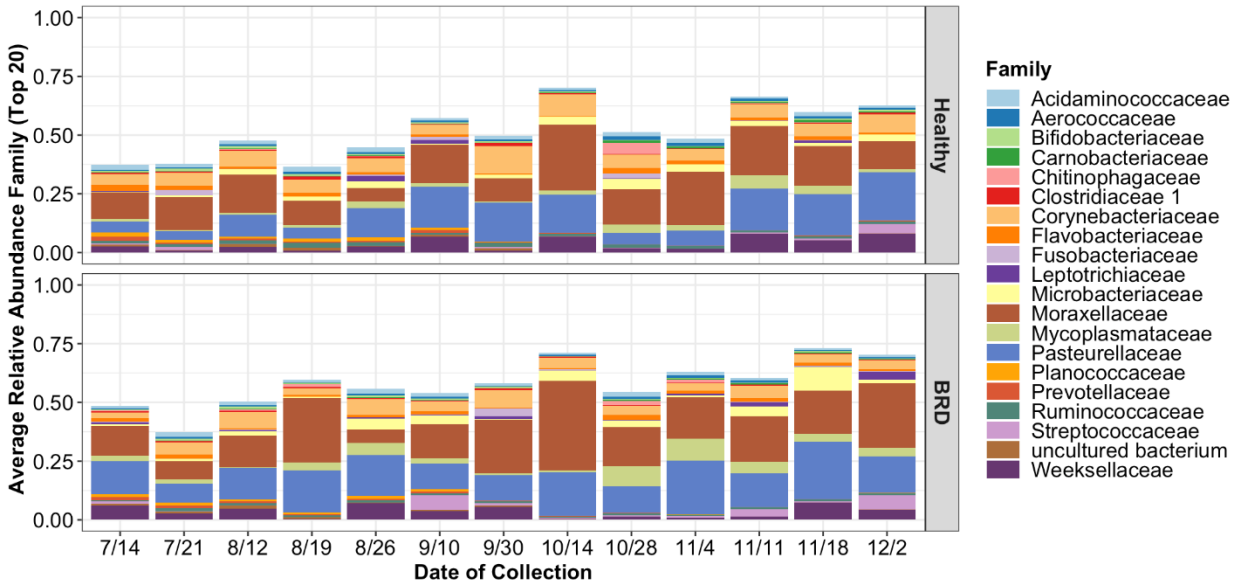

Figure A4. Stacked bar graph showing the average relative abundance of the top 20 family present in the nasal microbiome of BRD and healthy animals.

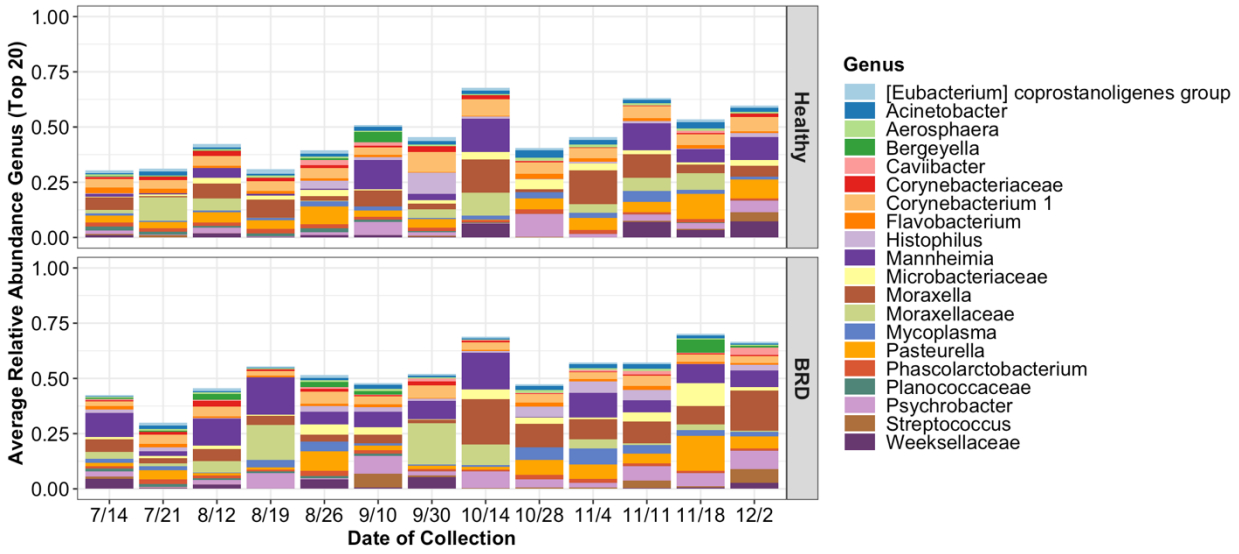

Figure A5. Stacked bar graph showing the average relative abundance of the top 20 genera present in the nasal microbiome of BRD and healthy animals.

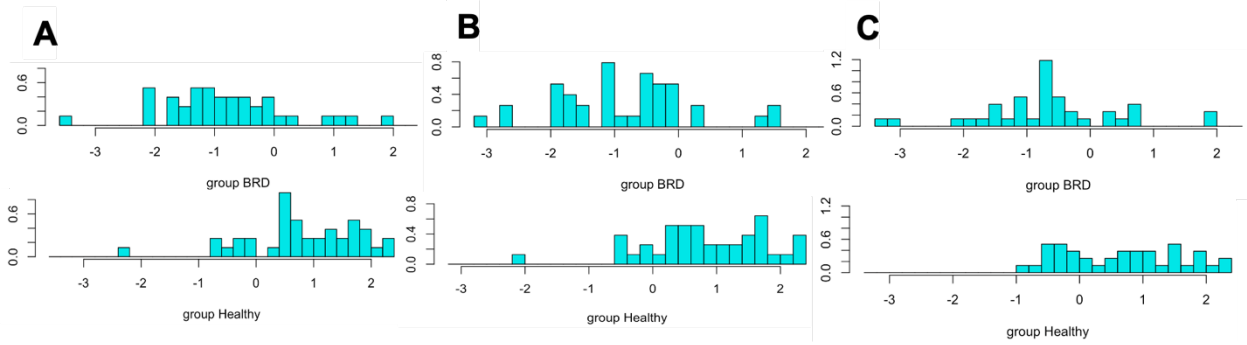

Figure A6. Visual representation of LDA models used to classify BRD and healthy animals using all samples (Analysis 1). Values lower than 0 represent animals classified as BRD, and values above 0 represent animals classified as healthy.

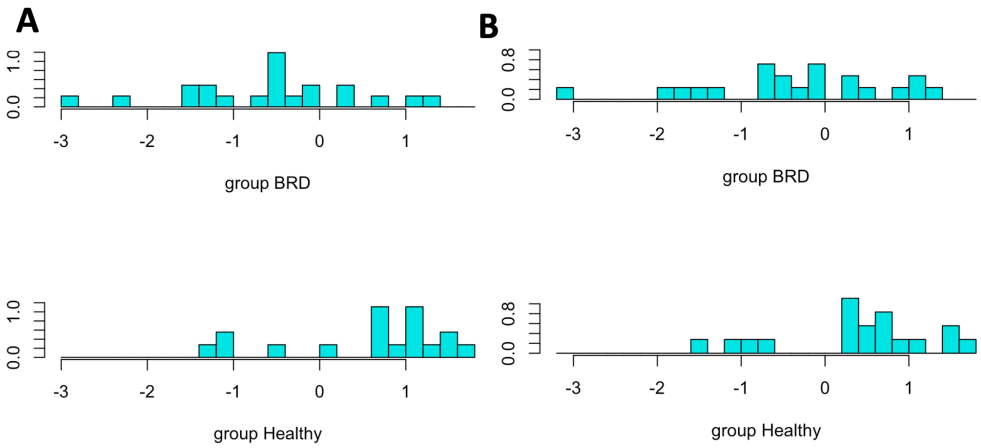

Figure A7. Visual representation of LDA models used to classify BRD and healthy animals using samples that tested positive to the four BRD-associated bacteria (Analysis 2). Values lower than 0 represent animals classified as BRD, and values above 0 represent animals classified as healthy.

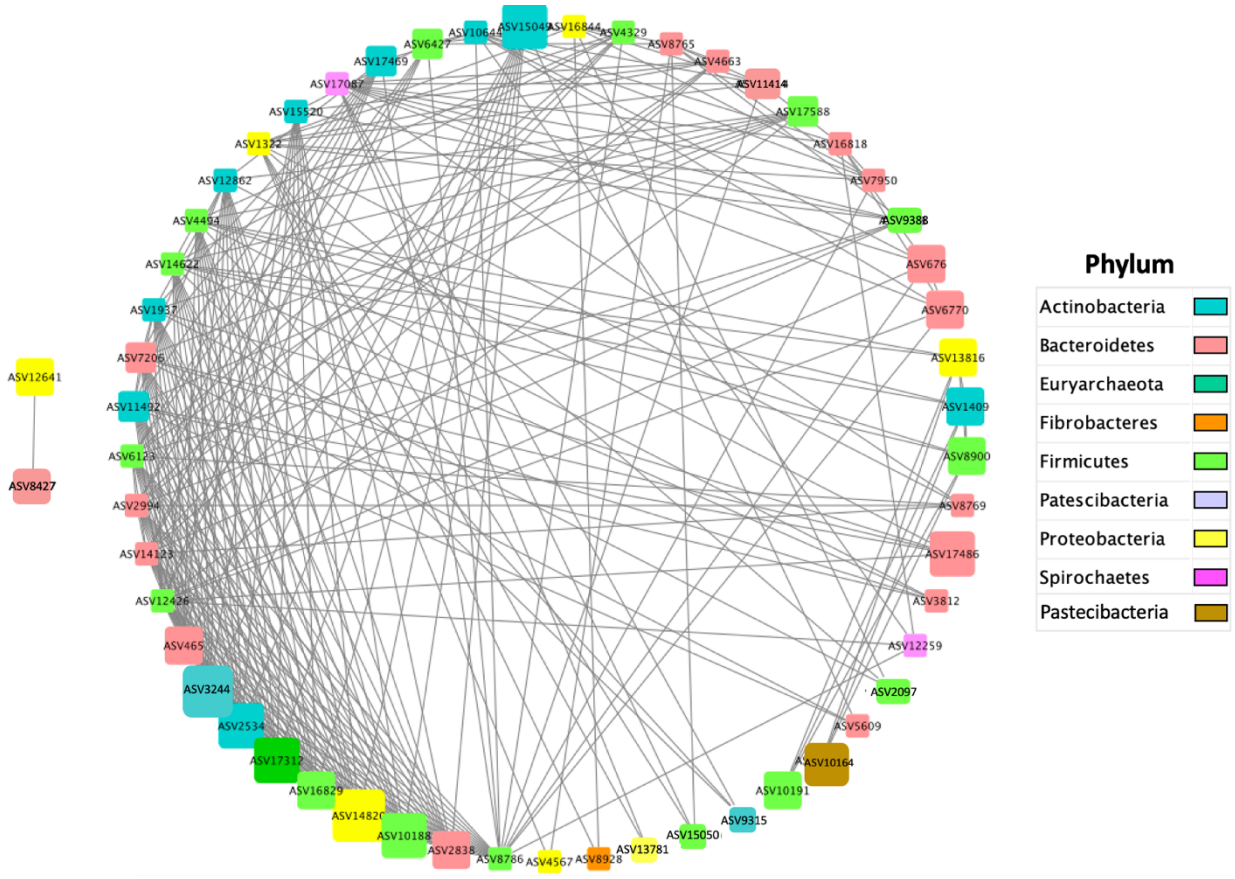

Figure A8. Visual representation of the co-occurrence analysis showing the abundant positive bacterial ASV pair combinations in the healthy group (n=74) with a probability of occurrence in the same sample greater than 0.9. Box sizes represent the total samples where the co-occurrence was observed (min value 69 and max value 73 samples). Color indicates the phylum level. Lines indicate how many associations each ASV had in the samples.

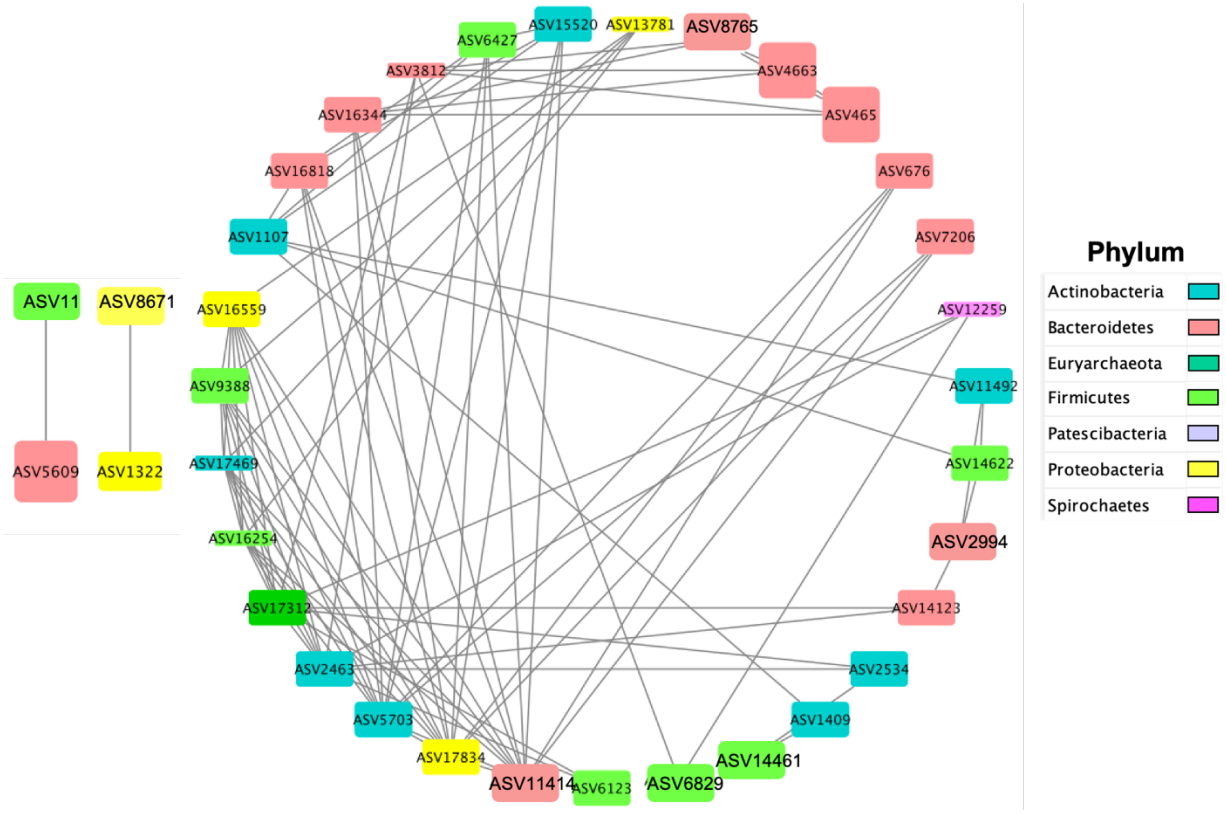

Figure A9. Visual representation of the co-occurrence analysis showing the abundant positive bacterial ASV pair combinations in the BRD group (n=57) with a probability of occurrence in the same sample greater than 0.9. Box sizes represent the total samples where the co-occurrence was observed (min value 54 and max value 56 samples). Color indicates the phylum level. Lines indicate how many associations each ASV had in the samples.

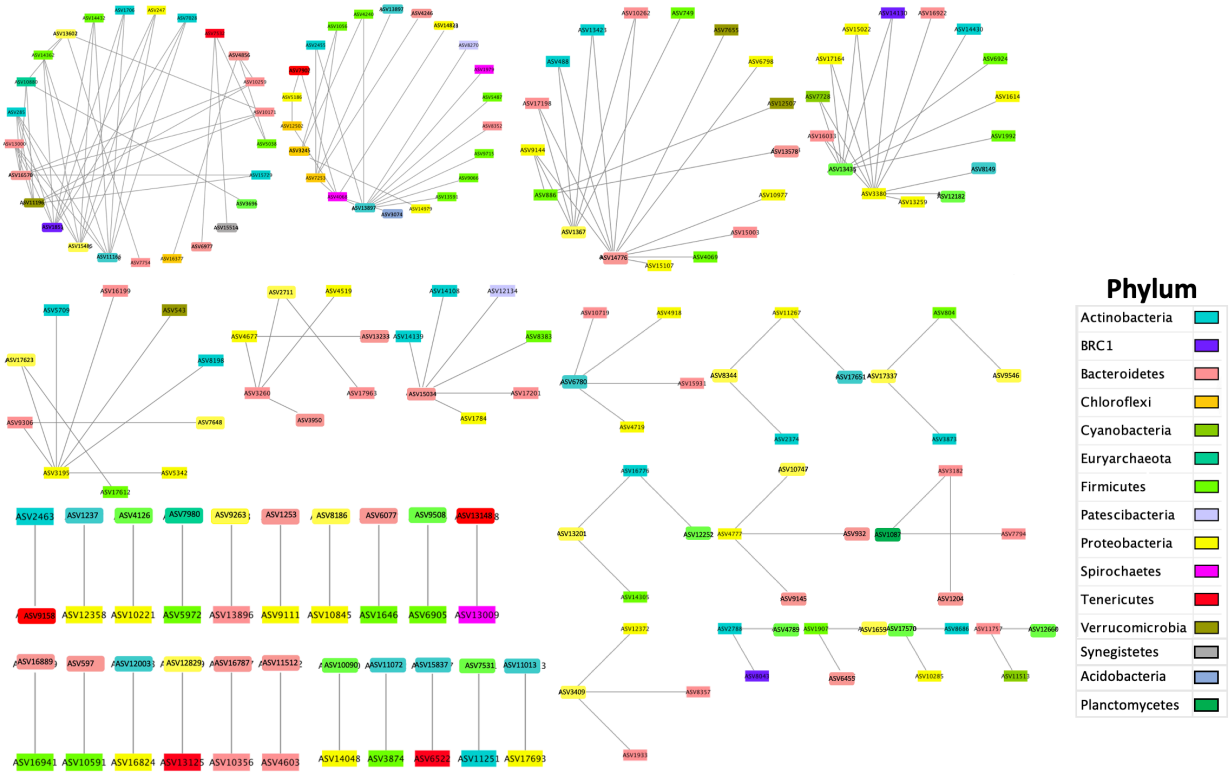

Figure A10. Visual representation of the co-occurrence analysis showing the negative ASV pair combinations in the healthy group (n=74) with a probability of occurrence in the same sample less than 0.05. Color indicates the phylum level. Lines indicate how many associations each ASV had in the samples.

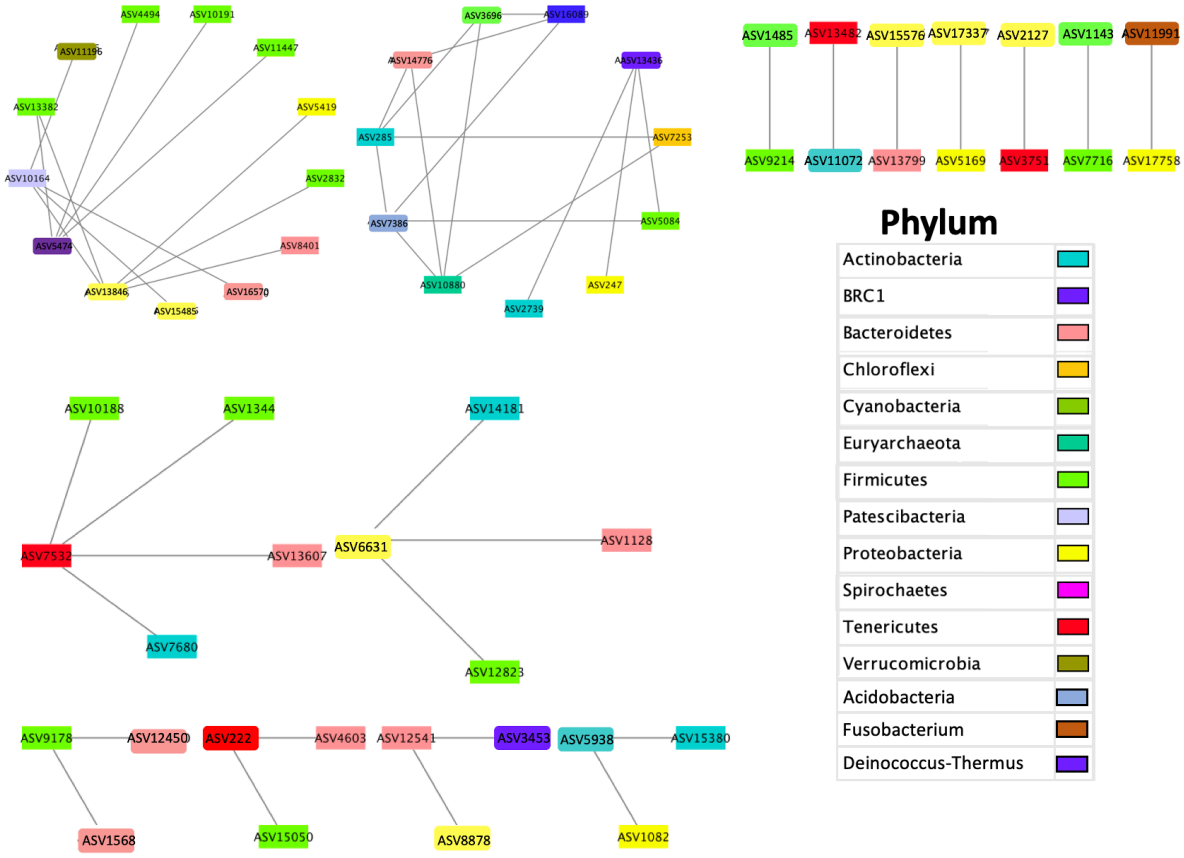

Figure A11. Visual representation of the co-occurrence analysis showing the negative ASV pair combinations in the BRD group (n=57) with a probability of occurrence in the same sample less than 0.05. Color indicates the phylum level. Lines indicate how many associations each ASV had in the samples.

Table A1. Sample average distance to the centroids of the BRD and healthy groups.

Beta diversity metric	**BRD**	**Healthy**	***P* Value**
Distance (Bray-Curtis)	0.5001	0.4743	0.058
Distance (Weighted UniFrac)	0.3369	0.3099	0.085

Table A2. Linear Discriminant equation including all samples (n=129) used to discriminate BRD and healthy animals.

Linear Discriminant equation including all samples (n=129) used to discriminate BRD and healthy animals	
Model 1	y=(-0.69415309*temp)+ (-1.46355431*Age)+(-0.32076803*Mb)+ (-0.09199251*Pm)+(1.08364767*Hs)+(-0.09380103*Mh)
Model 2	y=(-0.61621605*temp)+(-0.26230464*Mb)+ (0.05039419*Pm)+(0.90808658*Hs)+(-0.15137936*Mh)+(-0.27842520*16S rRNA)
Model 3	y=(-0.2790393*Mb)+(0.1654951*Pm)+ (1.0097024*Hs)+ (-0.2494981*Mh)+(-0.3915291*16S rRNA)

Temp = Rectal temperature
Mb= Mycoplasma bovis gene copy number (log10)
Pm= Pasteurella multocida gene copy number (log10)
Hs= Histophilus somni gene copy number (log10)
Mh= Mannheimina haemolytica gene copy number (log10)
16S rRNA= 16S rRNA gene copy number (log10)

Table A3. Linear Discriminant equation including only the samples that tested positive for the BRD-associated pathogens (n=66) used to discriminate BRD and healthy animals.

Linear Discriminant equation including only samples that tested positive for the BRD-associated pathogen (n=66) used to discriminate BRD and healthy animals.	
Model 1	y=(-0.5782106*16S rRNA) + (0.3368869*Pm) + (-3.9012453*Hs) + (0.2489542*Mh) + (-0.4398731*Mb) + (-0.7233586*Rectal temperature)
Model 2	y=(-0.7269952*Mb)+(0.9415095*Pm)+(-0.4415002*16S)

Mb= Mycoplasma bovis gene copy number (log10)
Pm= Pasteurella multocida gene copy number (log10)
Hs= Histophilus somni gene copy number (log10)
Mh= Mannheimina haemolytica gene copy number (log10)
16S rRNA= 16S rRNA gene copy number (log10)

Table A4. Summary of the Linea Discriminant Analysis showing the variation in the model misclassification error, percent of true positive (TP) and true negative (TN) identified and the sensitivity and specificity of the model when a predictor is removed from the analysis.

Possible combinations	**Predictor not included**	**Misclassification error**	**TP(%)**	**TN(%)**	**Sens (%)**	**Spec (%)**
Temperature, Age, *M. bovis, P. multocida, H. somni, M. haemolytica*, 16S rRNA	NA	0.30769923	23.1	46.2	57	77
Age, *M. bovis, P. multocida, H. somni, M. haemolytica*, 16S rRNA	Temperature	0.30769923	26.9	42.3	56	81
Temperature, *M. bovis, P. multocida, H. somni, M. haemolytica*, 16S rRNA	Age	0.30769923	23.1	46.2	57	77
Temperature, Age, *M. bovis, H. somni, M. haemolytica*, 16S rRNA	*P. multocida*	0.30769923	23.1	46.2	57	77
Temperature, Age, *M. bovis, P. multocida, H. somni, M. haemolytica*	16S rRNA	0.2884615	23.1	48.1	60	78
Temperature, Age, *P. multocida, H. somni, M. haemolytica*, 16S rRNA	*M. bovis*	0.2884615	17.3	53.8	64	73
Temperature, Age, *M. bovis, P. multocida, M. haemolytica*, 16S rRNA	*H. somni*	0.2884615	23.1	48.1	60	78
Temperature, Age, *M. bovis, P. multocida, H. somni,* 16S rRNA	*M. haemolytica*	0.2884615	23.1	48.1	60	78

Table A5. Co-occurrence analysis with positive ASV pair combinations between *Mycoplasma bovis* and member of the nasal microbiome present in healthy animals.

ASV_1	**ASV_2**	**obs_sites**	**probability**
Mycoplasma bovis	uncultured *Prevotellaceae* UCG-003	30	0.373
Mycoplasma bovis	*Prevotella* 9	30	0.373
Mycoplasma bovis	uncultured *Gulosibacter*	30	0.367
Mycoplasma bovis	uncultured *Sphingobacteriaceae*	30	0.362
Mycoplasma bovis	uncultured AKAU3644	30	0.362
Mycoplasma bovis	uncultured *Succinivibrio* sp.	29	0.351
Mycoplasma bovis	uncultured *Rikenellaceae* RC9 gut group	29	0.351
Mycoplasma bovis	*Arthrobacter*	29	0.351
Mycoplasma bovis	uncultured W5053	29	0.34
Mycoplasma bovis	uncultured *Fermentimonas*	30	0.34
Mycoplasma bovis	*Sphingomonadaceae*	29	0.34
Mycoplasma bovis	uncultured *Sphingobacterium*	28	0.334
Mycoplasma bovis	*Sanguibacter*	29	0.334
Mycoplasma bovis	*Aerococcus*	28	0.334
Mycoplasma bovis	uncultured *Bacteroides*	28	0.329
Mycoplasma bovis	unidentified *Oceanobacter*	29	0.323
Mycoplasma bovis	unidentified *Sphingobacterium*	27	0.318
Mycoplasma bovis	unidentified *Fastidiosipila*	28	0.318
Mycoplasma bovis	uncultured *Timonella*	27	0.318
Mycoplasma bovis	*Prevotellaceae U*CG-003	29	0.318
Mycoplasma bovis	*Pseudomonas saudiphocaensis*	27	0.312
Mycoplasma bovis	*Pseudomonas pertucinogena*	28	0.307
Mycoplasma bovis	*Methylophaga*	27	0.301
Mycoplasma bovis	*Leucobacter*	26	0.301
Mycoplasma bovis	*Fluviicola*	27	0.296
Mycoplasma bovis	uncultured *Ruminobacter*	27	0.29
Mycoplasma bovis	*Acholeplasma*	28	0.29
Mycoplasma bovis	uncultured *Tissierella*	27	0.285
Mycoplasma bovis	*Staphylococcus*	26	0.285
Mycoplasma bovis	*Ochrobactrum*	26	0.285
Mycoplasma bovis	uncultured *Muribaculaceae*	25	0.279
Mycoplasma bovis	*Myroides*	25	0.268
Mycoplasma bovis	*Azoarcus*	24	0.268
Mycoplasma bovis	*Cellvibrionaceae*	24	0.263
Mycoplasma bovis	uncultured *Proteiniphilum*	24	0.257
Mycoplasma bovis	uncultured *Prevotellaceae* UCG-003	24	0.257

Mycoplasma bovis	uncultured *Bacteroides* sp.	23	0.257
Mycoplasma bovis	*Acinetobacter*	23	0.257
Mycoplasma bovis	*Christensenellaceae* R-7 group	24	0.247
Mycoplasma bovis	uncultured *Succinivibrio* sp.	22	0.241
Mycoplasma bovis	*Stenotrophomonas*	22	0.241
Mycoplasma bovis	*Filobacterium*	22	0.236
Mycoplasma bovis	*Corynebacterium* 1	23	0.236
Mycoplasma bovis	uncultured *Fastidiosipila*	21	0.23
Mycoplasma bovis	uncultured *Proteiniphilum*	21	0.225
Mycoplasma bovis	uncultured *Coprococcus* 3	21	0.225
Mycoplasma bovis	uncultured *Carnobacteriaceae*	22	0.225
Mycoplasma bovis	*Marinococcus* sp. GSP31	21	0.225
Mycoplasma bovis	uncultured *Pseudoramibacter*	21	0.208
Mycoplasma bovis	*Lactococcus*	21	0.208
Mycoplasma bovis	*Treponema* 2	19	0.203
Mycoplasma bovis	*Sphingobacterium*	19	0.203
Mycoplasma bovis	uncultured *Ruminococcaceae*	20	0.197
Mycoplasma bovis	uncultured *Aerococcaceae*	20	0.197
Mycoplasma bovis	*Dysgonomonas*	21	0.197
Mycoplasma bovis	uncultured *Bacteroidales*	21	0.192
Mycoplasma bovis	*Erysipelothrix*	19	0.192
Mycoplasma bovis	*Ureaplasma diversum*	21	0.186
Mycoplasma bovis	*Carnobacteriaceae*	18	0.186
Mycoplasma bovis	*Acetobacter*	19	0.186
Mycoplasma bovis	uncultured *Salibacter* sp.	19	0.181
Mycoplasma bovis	*Facklamia*	19	0.181
Mycoplasma bovis	uncultured *Ruminococcaceae*	19	0.175
Mycoplasma bovis	uncultured *Thiopseudomonas*	19	0.17
Mycoplasma bovis	uncultured *Atopostipes*	18	0.17
Mycoplasma bovis	*Planococcaceae*	20	0.17
Mycoplasma bovis	*Leucobacter*	18	0.17
Mycoplasma bovis	uncultured *Akkermansia*	19	0.164
Mycoplasma bovis	*Pseudomonas*	17	0.159
Mycoplasma bovis	*Flavonifractor*	18	0.159
Mycoplasma bovis	uncultured *Peptococcaceae*	16	0.153
Mycoplasma bovis	*Corynebacterium*	17	0.153
Mycoplasma bovis	*Bacteroidetes bacterium* ADurb.Bin217	16	0.153
Mycoplasma bovis	uncultured *Salibacter* sp.	15	0.148
Mycoplasma bovis	uncultured *Atopostipes*	16	0.142

Mycoplasma bovis	*Lachnospiraceae*	15	0.137
Mycoplasma bovis	*Acholeplasma*	14	0.137
Mycoplasma bovis	uncultured *Leucobacter*	15	0.131
Mycoplasma bovis	uncultured *Halomonas*	14	0.131
Mycoplasma bovis	uncultured compost *Brumimicrobium*	15	0.131
Mycoplasma bovis	*Peptoniphilus*	15	0.126
Mycoplasma bovis	*Mycoplasma*	17	0.126
Mycoplasma bovis	*Treponema succinifaciens* DSM 2489	12	0.104
Mycoplasma bovis	*Mycoplasma*	11	0.099
Mycoplasma bovis	uncultured *Proteiniphilum*	10	0.088
Mycoplasma bovis	uncultured *Fastidiosipila*	10	0.088
Mycoplasma bovis	uncultured *Tissierella* sp.	11	0.082
Mycoplasma bovis	uncultured *Fluviicola*	8	0.066
Mycoplasma bovis	*Erysipelothrix*	10	0.066
Mycoplasma bovis	*Staphylococcus*	8	0.06
Mycoplasma bovis	uncultured *Alloprevotella*	9	0.055
Mycoplasma bovis	uncultured *Treponema* 2	6	0.044
Mycoplasma bovis	*Mycoplasma* sp. LR5794	5	0.033
Mycoplasma bovis	*Mycoplasma arginini*	6	0.033
Mycoplasma bovis	*Moheibacter stercoris*	5	0.033
Mycoplasma bovis	uncultured *Alysiella*	5	0.027

Table A6. Co-occurrence analysis with negative ASV pair combinations between *Mycoplasma bovis* and member of the nasal microbiome present in healthy animals.

ASV_1	***ASV_2***	**obs_sites**	**probability**
Mycoplasma bovis	uncultured *Clostridiales*	26	0.383
Mycoplasma bovis	*Streptococcus*	19	0.318
Mycoplasma bovis	uncultured *Thermomicrobiales*	18	0.301
Mycoplasma bovis	uncultured *Arenimonas*	17	0.285
Mycoplasma bovis	*Megasphaera elsdenii*	16	0.268
Mycoplasma bovis	uncultured *Saprospiraceae*	14	0.268
Mycoplasma bovis	*Ruminococcaceae* UCG-005	15	0.263
Mycoplasma bovis	*Hydrogenophaga*	14	0.257
Mycoplasma bovis	uncultured *Persicitalea*	14	0.252
Mycoplasma bovis	uncultured *Olsenella*	13	0.252
Mycoplasma bovis	human gut metagenome- *Clostridium sensu stricto* 1	14	0.247
Mycoplasma bovis	*Streptococcus henryi*	13	0.247

Mycoplasma bovis	*Arenimonas*	11	0.236
Mycoplasma bovis	uncultured compost *Pusillimonas*	12	0.225
Mycoplasma bovis	*Lachnospiraceae bacterium*	11	0.214
Mycoplasma bovis	uncultured *Altererythrobacter*	11	0.203
Mycoplasma bovis	*Agathobacter*	11	0.203
Mycoplasma bovis	*Gordonia*	11	0.203
Mycoplasma bovis	uncultured *Advenella*	11	0.203
Mycoplasma bovis	uncultured *Planctomyces* sp.	11	0.203
Mycoplasma bovis	uncultured *Verrucomicrobium*	10	0.197
Mycoplasma bovis	uncultured *Faecalibacterium* sp.	9	0.197
Mycoplasma bovis	*Rhodobacteraceae*	10	0.192
Mycoplasma bovis	*Clostridium sensu stricto* 1	10	0.192
Mycoplasma bovis	uncultured *Mogibacterium*	10	0.192
Mycoplasma bovis	uncultured *Roseimaritima*	9	0.192
Mycoplasma bovis	*Fusicatenibacter*	9	0.186
Mycoplasma bovis	uncultured *Taibaiella*	9	0.181
Mycoplasma bovis	*Enterococcus*	9	0.181
Mycoplasma bovis	*Aquiflexum*	7	0.181
Mycoplasma bovis	uncultured *Thermus/Deinococcus* group bacterium	9	0.175
Mycoplasma bovis	uncultured *Ornithinimicrobium*	9	0.175
Mycoplasma bovis	uncultured *Chitinophagales*	9	0.175
Mycoplasma bovis	*Paraburkholderia tropica*	8	0.17
Mycoplasma bovis	*Paraburkholderia tropica*	7	0.159
Mycoplasma bovis	*Blautia*	6	0.153
Mycoplasma bovis	uncultured *Rhodobacteraceae*	7	0.148
Mycoplasma bovis	uncultured *Cellvibrio*	7	0.148
Mycoplasma bovis	uncultured *Syntrophococcus*	6	0.148
Mycoplasma bovis	*Rhodobacteraceae*	5	0.148
Mycoplasma bovis	*Absconditabacteriales* (SR1)	6	0.142
Mycoplasma bovis	uncultured *anaerobic bacterium*	6	0.137
Mycoplasma bovis	*Cellulomonas*	6	0.137
Mycoplasma bovis	*Staphylococcus lentus*	6	0.137
Mycoplasma bovis	*Flavobacterium* sp. enrichment culture clone SA_NR2_1	6	0.131
Mycoplasma bovis	*Petrimonas*	6	0.131
Mycoplasma bovis	uncultured *Verrucomicrobium*	6	0.131
Mycoplasma bovis	*Dietzia*	6	0.131
Mycoplasma bovis	*Veillonella magna*	6	0.131

Mycoplasma bovis	*Fermentimonas*	6	0.131
Mycoplasma bovis	uncultured *Prevotellaceae* UCG-001	5	0.131
Mycoplasma bovis	*Roseimaritima*	4	0.126
Mycoplasma bovis	*Pseudomonas thermotolerans* J53	4	0.126
Mycoplasma bovis	*Stenotrophomonas*	5	0.121
Mycoplasma bovis	*Hyaloperonospora arabidopsidis*	5	0.121
Mycoplasma bovis	hydrothermal vent metagenome	5	0.121
Mycoplasma bovis	*Christensenellaceae* R-7 group	4	0.115
Mycoplasma bovis	*Leucobacter*	4	0.115
Mycoplasma bovis	JG30-KF-CM45	4	0.115
Mycoplasma bovis	*Microbacteriaceae*	4	0.104
Mycoplasma bovis	uncultured *Chelatococcus* sp.	4	0.104
Mycoplasma bovis	*Christensenellaceae* R-7 group	4	0.104
Mycoplasma bovis	*Flavobacterium*	3	0.099
Mycoplasma bovis	uncultured *Proteiniborus*	3	0.099
Mycoplasma bovis	*Cellulomonas*	3	0.088
Mycoplasma bovis	uncultured *Proteiniborus*	2	0.082
Mycoplasma bovis	*Paracoccus*	2	0.071
Mycoplasma bovis	uncultured *Sanguibacter*	2	0.071
Mycoplasma bovis	*Leucobacter*	1	0.071
Mycoplasma bovis	uncultured *Glutamicibacter*	1	0.06
Mycoplasma bovis	*Sphingobacterium*	1	0.06
Mycoplasma bovis	*Sphingobacterium*	1	0.055
Mycoplasma bovis	uncultured *Acidobacteria*	0	0.044
Mycoplasma bovis	*Luteimonas*	0	0.033

Table A7. Co-occurrence analysis with positive ASV pair combinations between *Mycoplasma bovis* and member of the nasal microbiome present in BRD animals.

ASV_1	**ASV_2**	**obs_sites**	**probability**
Mycoplasma bovis	*Filobacterium*	35	0.547
Mycoplasma bovis	*Clostridiaceae* SK061	34	0.547
Mycoplasma bovis	uncultured *Prevotella* 2	34	0.547
Mycoplasma bovis	uncultured AKAU3644	34	0.524
Mycoplasma bovis	*Aerococcus*	31	0.478
Mycoplasma bovis	uncultured *Succinivibrio*	30	0.456
Mycoplasma bovis	uncultured *Tissierella*	28	0.433
Mycoplasma bovis	*Ureaplasma diversum*	29	0.433
Mycoplasma bovis	uncultured *Anaerosporobacter*	28	0.421

Mycoplasma bovis	*Bacteroides*	27	0.41
Mycoplasma bovis	uncultured *Succinivibrio* sp.	23	0.342
Mycoplasma bovis	*Acetitomaculum*	22	0.319
Mycoplasma bovis	uncultured *Marvinbryantia*	22	0.307
Mycoplasma bovis	*Mycoplasma*	22	0.285
Mycoplasma bovis	*Lactococcus*	20	0.273
Mycoplasma bovis	uncultured *Lachnospiraceae* UCG-007	19	0.251
Mycoplasma bovis	*Mycoplasma*	20	0.251
Mycoplasma bovis	*Lachnospiraceae*	20	0.251
Mycoplasma bovis	*Mycoplasma*	21	0.251
Mycoplasma bovis	*Ochrobactrum*	17	0.239
Mycoplasma bovis	uncultured *Bacteroidales*	17	0.239
Mycoplasma bovis	*Aliidiomarina*	16	0.216
Mycoplasma bovis	*Mycoplasma*	16	0.216
Mycoplasma bovis	*Mycoplasma alkalescens* 14918	17	0.205
Mycoplasma bovis	*Moraxellaceae*	13	0.171
Mycoplasma bovis	*Alloprevotella*	13	0.159
Mycoplasma bovis	*Christensenellaceae* R-7 group	11	0.137
Mycoplasma bovis	*Methanobrevibacter*	11	0.137
Mycoplasma bovis	*Mycoplasma arginini*	11	0.137
Mycoplasma bovis	*Prevotella* 9	10	0.125
Mycoplasma bovis	*Paraburkholderia tropica*	10	0.125
Mycoplasma bovis	uncultured rumen *Prevotellaceae* UCG-001	10	0.114
Mycoplasma bovis	*Ruminococcus* sp. YE281	10	0.114
Mycoplasma bovis	uncultured *Prevotella* 9	10	0.114
Mycoplasma bovis	*Brachybacterium*	9	0.102
Mycoplasma bovis	uncultured *Coprococcus* 3	8	0.091
Mycoplasma bovis	*Idiomarina*	8	0.091
Mycoplasma bovis	uncultured *Jeotgalicoccus*	8	0.091
Mycoplasma bovis	*Lachnospiraceae*	7	0.08
Mycoplasma bovis	*Pasteurellaceae*	7	0.08
Mycoplasma bovis	*Lachnospiraceae*	7	0.08
Mycoplasma bovis	*Staphylococcus*	7	0.08
Mycoplasma bovis	uncultured *Muribaculaceae*	4	0.025

Table A8. Co-occurrence analysis with negative ASV pair combinations between *Mycoplasma bovis* and member of the nasal microbiome present in BRD animals.

ASV_1	**ASV_2**	**obs_sites**	**probability**
Mycoplasma bovis	uncultured *Firmicutes*	27	0.524
Mycoplasma bovis	*Arthrobacter*	24	0.478
Mycoplasma bovis	uncultured *Oligella*	22	0.456
Mycoplasma bovis	uncultured *Porphyromonadaceae*	21	0.433
Mycoplasma bovis	*Intrasporangiaceae*	20	0.433
Mycoplasma bovis	uncultured *Clostridiales*	20	0.421
Mycoplasma bovis	*Hydrogenophaga*	20	0.41
Mycoplasma bovis	uncultured *Bacteroidales*	19	0.399
Mycoplasma bovis	*Sphingo* sp. 773B2_12ER2A	18	0.387
Mycoplasma bovis	*Tissierella*	18	0.376
Mycoplasma bovis	*Corynebacterium*	16	0.376
Mycoplasma bovis	*Christensenellaceae* R-7 group	18	0.376
Mycoplasma bovis	*Chishuiella* sp. YIM 102668	15	0.376
Mycoplasma bovis	*Burkholderiaceae*	16	0.364
Mycoplasma bovis	*Actinomyces*	16	0.364
Mycoplasma bovis	*Acinetobacter*	16	0.364
Mycoplasma bovis	uncultured compost *Pusillimonas*	16	0.353
Mycoplasma bovis	*Rhodobacteraceae*	16	0.353
Mycoplasma bovis	uncultured *Proteiniclasticum*	15	0.342
Mycoplasma bovis	uncultured *Leucobacter sp.*	16	0.342
Mycoplasma bovis	uncultured *Clostridium* sp.	16	0.342
Mycoplasma bovis	*Planococcaceae*	14	0.342
Mycoplasma bovis	*Methanobrevibacter*	16	0.342
Mycoplasma bovis	*Devosia*	16	0.342
Mycoplasma bovis	*Aquiflexum*	15	0.33
Mycoplasma bovis	uncultured *Advenella*	14	0.319
Mycoplasma bovis	*Fluviicola*	14	0.319
Mycoplasma bovis	*Streptococcus henryi*	13	0.307
Mycoplasma bovis	*Brevibacterium*	14	0.307
Mycoplasma bovis	uncultured *Truepera*	13	0.296
Mycoplasma bovis	uncultured *Treponema* 2	13	0.296
Mycoplasma bovis	uncultured *Moheibacter*	13	0.296
Mycoplasma bovis	*Psychrobacter maritimus*	13	0.296
Mycoplasma bovis	*Filobacterium*	13	0.296
Mycoplasma bovis	*Chryseobacterium* sp. A5	13	0.296
Mycoplasma bovis	uncultured *Alistipes*	12	0.285

Mycoplasma bovis	*Prevotellaceae*	12	0.285
Mycoplasma bovis	*Peptoniphilus*	11	0.285
Mycoplasma bovis	*Gordonia*	12	0.285
Mycoplasma bovis	uncultured *Chitinophagales*	12	0.273
Mycoplasma bovis	*Stenotrophomonas*	12	0.273
Mycoplasma bovis	*Rhodocyclaceae*	11	0.273
Mycoplasma bovis	Family XI	12	0.273
Mycoplasma bovis	uncultured *Verrucomicrobiaceae*	11	0.262
Mycoplasma bovis	uncultured *Ornithinimicrobium*	11	0.262
Mycoplasma bovis	uncultured *Chitinophagales*	11	0.262
Mycoplasma bovis	*Flavobacteriales*	11	0.262
Mycoplasma bovis	*Bifidobacterium*	11	0.262
Mycoplasma bovis	uncultured *Oceanobacter*	10	0.251
Mycoplasma bovis	*Leucobacter*	10	0.251
Mycoplasma bovis	uncultured *Paludibacter* sp.	10	0.239
Mycoplasma bovis	uncultured *Christensenellaceae* R-7 group	10	0.239
Mycoplasma bovis	uncultured *AKAU3644*	8	0.239
Mycoplasma bovis	uncultured *Petrimonas*	8	0.228
Mycoplasma bovis	*Rhodobacteraceae*	9	0.228
Mycoplasma bovis	*Clostridium sensu stricto* 1	9	0.228
Mycoplasma bovis	uncultured *Thermus/Deinococcus* group	9	0.216
Mycoplasma bovis	uncultured rumen *Mogibacterium*	9	0.216
Mycoplasma bovis	uncultured gamma proteo *Alcanivorax*	9	0.216
Mycoplasma bovis	uncultured *Flavobacterium*	8	0.216
Mycoplasma bovis	uncultured *Camelimonas*	9	0.216
Mycoplasma bovis	uncultured BRC1	9	0.216
Mycoplasma bovis	uncultured *Atopostipes*	9	0.216
Mycoplasma bovis	*Leucobacter*	9	0.216
Mycoplasma bovis	*Leadbetterella byssophila* DSM 17132	9	0.216
Mycoplasma bovis	*Brachybacterium*	8	0.216
Mycoplasma bovis	uncultured *Taibaiella*	8	0.205
Mycoplasma bovis	uncultured *Gammaproteobacteria*	8	0.205
Mycoplasma bovis	uncultured *Flaviflexus*	8	0.205
Mycoplasma bovis	*Mannheimia*	8	0.205
Mycoplasma bovis	*BD1-7* clade	8	0.205
Mycoplasma bovis	uncultured *Prevotella* 1	6	0.194
Mycoplasma bovis	uncultured *Paludibacteraceae*	6	0.194
Mycoplasma bovis	uncultured *Microtrichaceae*	7	0.194
Mycoplasma bovis	uncultured *Cellvibrio*	7	0.194

Mycoplasma bovis	*Christensenellaceae* R-7 group	6	0.194
Mycoplasma bovis	uncultured *Mogibacterium*	6	0.182
Mycoplasma bovis	uncultured *Clostridium* sp.	7	0.182
Mycoplasma bovis	uncultured *Helcococcus*	6	0.171
Mycoplasma bovis	*Sphingobacterium* sp. 35	6	0.171
Mycoplasma bovis	*Nocardioides*	6	0.171
Mycoplasma bovis	hydrothermal vent metagenome	6	0.171
Mycoplasma bovis	uncultured *Leucobacter*	5	0.159
Mycoplasma bovis	uncultured *Burkholderiaceae*	6	0.159
Mycoplasma bovis	uncultured *Bogoriella*	6	0.159
Mycoplasma bovis	*Christensenellaceae* R-7 group	5	0.148
Mycoplasma bovis	uncultured *Bacteroidetes*	4	0.137
Mycoplasma bovis	*Rhodobacteraceae*	4	0.125
Mycoplasma bovis	*Burkholderiaceae*	4	0.125
Mycoplasma bovis	*Ruminococcaceae*	3	0.102
Mycoplasma bovis	uncultured *Clostridiales*	2	0.08
Mycoplasma bovis	*Paracoccus*	2	0.08
Mycoplasma bovis	*Microbacteriaceae*	2	0.08
Mycoplasma bovis	*Desulfovibrio piger*	1	0.057
Mycoplasma bovis	uncultured *Muribaculaceae*	0	0.043
Mycoplasma bovis	*Lachnospiraceae*	0	0.04

APPENDIX B

Table B1. Metadata containing the cattle's sample information used in this study.

ID	Cattle ID	BRD	Date/Collection	Ave.temp	Temp	Age	PenCode
Dairy1	Black 151	Healthy	8/12/20	43.71	103.2	0.88	10
Dairy2	Black 17	Healthy	7/21/20	74	104.4	0.82	11
Dairy3	Black 192	BRD	7/14/20	72.71	104.3	0.8	10
Dairy4	Black 220	Healthy	7/14/20	72.71	103.7	0.8	10
Dairy5	Black 25	Healthy	7/21/20	74	104.7	0.82	11
Dairy6	Black 337	Healthy	7/14/20	72.71	102.7	0.76	8
Dairy7	Black 340	BRD	7/14/20	72.71	105.6	0.76	8
Dairy8	Black 350	Healthy	7/14/20	72.71	104	0.76	10
Dairy9	Black 377	Healthy	7/14/20	72.71	102.9	0.76	8
Dairy10	Black 387	Healthy	7/21/20	74	103.3	0.78	7
Dairy11	Black 388	BRD	7/21/20	74		0.78	4
Dairy12	Black 454	BRD	7/21/20	74	102.7	0.78	7
Dairy13	Black 458	BRD	7/21/20	74	102.9	0.78	7
Dairy14	Black 492	BRD	7/14/20	72.71	103.8	0.76	8
Dairy15	Black 566	BRD	7/21/20	74	103.3	0.68	4
Dairy16	Black 603	BRD	9/30/20	54.88	104.6	0.88	4
Dairy17	Black 659	Healthy	7/14/20	72.71	103	0.71	6
Dairy18	Black 661	BRD	7/14/20	72.71	105.2	0.71	6
Dairy19	Black 665	Healthy	7/14/20	72.71	102.5	0.71	6
Dairy20	Black 709	Healthy	8/12/20	43.71	102.7	0.74	10
Dairy21	Black 769	Healthy	12/2/20	33.04	103.7	1.05	10
Dairy22	Black 884	Healthy	8/26/20	80.29	105	0.74	6
Dairy23	Black 899	Healthy	9/10/20	62.08	102.9	0.78	8
Dairy24	Black 901	Healthy	8/19/20	66.25		0.8	5
Dairy25	Black 903	Healthy	8/19/20	66.25		0.8	6
Dairy26	Black 917	Healthy	9/10/20	62.08	104	0.78	8
Dairy27	Black 921	BRD	8/12/20	43.71	102.4	0.7	6
Dairy28	Black 969	BRD	8/19/20	66.25	104.4	0.8	6
Dairy29	Black 973	BRD	8/12/20	43.71	104.8	0.7	6
Dairy30	Black 985	Healthy	8/26/20	80.29	105.1	0.74	6
Dairy31	Black 986	Healthy	8/19/20	66.25	103.3	0.8	5
Dairy32	Black 989	Healthy	8/12/20	43.71	103.4	0.7	5
Dairy33	Blue 10	Healthy	11/18/20	40.08	103.7	0.67	5

Dairy34	Blue 102	Healthy	11/18/20	40.08	102.5	0.585	1
Dairy35	Blue 105	BRD	11/4/20	60.5	104.1	0.55	3
Dairy36	Blue 121	Healthy	11/18/20	40.08	103.6	0.585	3
Dairy37	Blue 122	Healthy	10/28/20	41.11	103.6	0.53	3
Dairy38	Blue 129	BRD	10/28/20	41.11	107.4	0.53	3
Dairy39	Blue 139	Healthy	11/18/20	40.08	103.3	0.585	3
Dairy40	Blue 167	Healthy	10/28/20	41.11	102.5	0.53	1
Dairy41	Blue 180	Healthy	11/11/20	40	102.9	0.57	2
Dairy42	Blue 181	BRD	10/28/20	41.11	104.9	0.53	1
Dairy43	Blue 183	Healthy	11/18/20	40.08	102.4	0.585	1
Dairy44	Blue 187	Healthy	11/4/20	60.5	102.7	0.55	3
Dairy45	Blue 192	Healthy	10/28/20	41.11	103.4	0.53	3
Dairy46	Blue 271	BRD	11/4/20	60.5	105.6	0.55	3
Dairy47	Blue 371	Healthy	10/28/20	41.11		0.45	0
Dairy48	Blue 373	BRD	10/28/20	41.11	106.2	0.45	0
Dairy49	Blue 374	BRD	10/28/20	41.11	106	0.45	0
Dairy50	Blue 383	BRD	11/4/20	60.5	104.6	0.47	0
Dairy51	Blue 384	Healthy	11/4/20	60.5	102.2	0.47	0
Dairy52	Blue 385	Healthy	11/4/20	60.5	103.4	0.47	0
Dairy53	Blue 386	BRD	11/4/20	60.5	104.1	0.47	0
Dairy54	Blue 415	BRD	12/2/20	33.04	103	0.54	1
Dairy55	Blue 416	Healthy	12/2/20	33.04	102.9	0.54	1
Dairy56	Blue 417	Healthy	12/2/20	33.04	103.5	0.54	1
Dairy57	Blue 43	Healthy	10/28/20	41.11	103.3	0.525	3
Dairy58	Blue 62	BRD	11/18/20	40.08	105.2	0.58	1
Dairy59	Blue 64	Healthy	11/18/20	40.08	102.5	0.58	3
Dairy60	Blue 8	BRD	11/11/20	40	106.5	0.56	4
Dairy61	Blue 86	Healthy	11/11/20	40	106.6	0.56	2
Dairy62	Green 905	Healthy	7/21/20	74	103.3	1.45	0
Dairy63	Green 918	Healthy	7/21/20	74	102.3	1.45	0
Dairy64	Orange 106	Healthy	8/12/20	43.71	103.1	0.7	5
Dairy65	Orange 11	BRD	8/12/20	43.71	106.3	0.7	5
Dairy66	Orange 110	Healthy	8/26/20	80.29	103.4	0.74	2
Dairy67	Orange 118	BRD	8/26/20	80.29	104.2	0.74	6
Dairy68	Orange 16	BRD	7/14/20	72.71	104.6	0.62	3
Dairy69	Orange 205	BRD	9/30/20	54.88	104.1	0.755	6
Dairy70	Orange 220	Healthy	8/26/20	80.29	102.7	0.61	3
Dairy71	Orange 234	Healthy	9/30/20	54.88	102.7	0.71	4
Dairy72	Orange 241	BRD	8/26/20	80.29	102.7	0.61	3

Dairy73	Orange 246	Healthy	9/30/20	54.88	103.1	0.71	6
Dairy74	Orange 248	BRD	9/10/20	62.08	103.6	0.65	4
Dairy75	Orange 253	Healthy	8/26/20	80.29	103.5	0.61	1
Dairy76	Orange 26	Healthy	7/14/20	72.71	102.5	0.62	3
Dairy77	Orange 262	BRD	8/26/20	80.29	105.2	0.61	2
Dairy78	Orange 282	BRD	8/26/20	80.29	102.9	0.61	3
Dairy79	Orange 293	BRD	9/10/20	62.08	103	0.65	4
Dairy80	Orange 311	Healthy	9/10/20	62.08	103.1	0.65	4
Dairy81	Orange 327	Healthy	9/30/20	54.88	102.9	0.71	6
Dairy82	Orange 361	Healthy	8/26/20	80.29	103.3	0.61	1
Dairy83	Orange 4	Healthy	7/14/20	72.71	102.6	0.62	3
Dairy84	Orange 425	Healthy	8/26/20	80.29	104.1	0.61	3
Dairy85	Orange 466	BRD	8/26/20	80.29	103.4	0.61	1
Dairy86	Orange 489	Healthy	8/26/20	80.29	103.4	0.61	2
Dairy87	Orange 505	BRD	9/30/20	54.88	103.6	0.645	3
Dairy88	Orange 519	BRD	7/21/20	74	104.7	0.45	0
Dairy89	Orange 534	Healthy	9/30/20	54.88	102.9	0.645	3
Dairy90	Orange 549	Healthy	9/30/20	54.88	103	0.645	3
Dairy91	Orange 577	Healthy	10/14/20	57.63	102.7	0.68	6
Dairy92	Orange 581	BRD	10/14/20	57.63	103.4	0.68	6
Dairy93	Orange 587	BRD	10/14/20	57.63	104.7	0.68	6
Dairy94	Orange 6	BRD	8/19/20	66.25	103.2	0.72	5
Dairy95	Orange 604	BRD	9/10/20	62.08	104.6	0.72	3
Dairy96	Orange 605	Healthy	9/10/20	62.08	103.4	0.72	3
Dairy97	Orange 606	Healthy	9/10/20	62.08	102.8	0.72	3
Dairy98	Orange 607	BRD	9/10/20	62.08	102.5	0.72	3
Dairy99	Orange 628	Healthy	10/14/20	57.63	103.3	0.68	6
Dairy100	Orange 63	BRD	9/10/20	62.08	103.2	0.78	8
Dairy101	Orange 634	BRD	9/30/20	54.88	106.5	0.645	3
Dairy102	Orange 641	Healthy	10/14/20	57.63	102.7	0.68	5
Dairy103	Orange 653	BRD	10/14/20	57.63	105.2	0.68	5
Dairy104	Orange 655	Healthy	9/30/20	54.88	102.9	0.645	2
Dairy105	Orange 682	Healthy	9/30/20	54.88	102.5	0.645	2
Dairy106	Orange 705	BRD	9/30/20	54.88	103.3	0.645	2
Dairy107	Orange 713	BRD	9/30/20	54.88	103.9	0.645	2
Dairy108	Orange 736	Healthy	10/14/20	57.63	105.4	0.68	5
Dairy109	Orange 747	Healthy	11/11/20	40	102.9	0.76	4
Dairy110	Orange 765	BRD	12/2/20	33.04	104	0.71	8
Dairy111	Orange 77	BRD	8/19/20	66.25	103.8	0.72	6

Dairy112	Orange 778	Healthy	11/11/20	40	103.8	0.65	6
Dairy113	Orange 782	BRD	11/18/20	40.08	104.8	0.67	5
Dairy114	Orange 789	BRD	11/11/20	40	104.4	0.65	6
Dairy115	Orange 83	Healthy	8/12/20	43.71	102.7	0.7	6
Dairy116	Orange 838	BRD	12/2/20	33.04	104.9	0.71	9
Dairy117	Orange 841	Healthy	11/11/20	40	103.2	0.65	4
Dairy118	Orange 847	Healthy	12/2/20	33.04	102.5	0.71	8
Dairy119	Orange 865	Healthy	11/18/20	40.08	102.9	0.67	5
Dairy120	Orange 88	Healthy	8/12/20	43.71	103.2	0.7	6
Dairy121	Orange 909	Healthy	11/4/20	60.5	102.9	0.63	5
Dairy122	Orange 919	Healthy	12/2/20	33.04	104.9	0.71	9
Dairy123	Orange 922	BRD	11/18/20	40.08	105.8	0.67	5
Dairy124	Orange 924	Healthy	11/11/20	40	104	0.65	6
Dairy125	Orange 939	BRD	12/2/20	33.04	103.8	0.71	8
Dairy126	Orange 941	Healthy	12/2/20	33.04	103	0.71	9
Dairy127	Orange 988	BRD	11/4/20	60.5	105.1	0.63	5
Dairy128	Orange 990	Healthy	12/2/20	33.04	102	0.71	8
Dairy129	Orange 994	BRD	11/11/20	40	104.8	0.65	6
Dairy130	Purple 234	BRD	12/2/20	33.04	102.7	0.33	10
Dairy131	Yellow 119	BRD	8/12/20	43.71	103.5	1.15	10
Dairy132	Yellow 591	BRD	8/12/20	43.71	103.3	1	10
Dairy133	Yellow 869	BRD	7/21/20	74	104	0.82	11

**BRD: indicates the cattle health status*
**Ave.Temp: Daily Average Temperature of the date when the nasal swab was collected*
**Temp: Rectal temperature of the cattle that was sampled*
**Age: Age of the cattle that was sampled*
**PenCode: Pen number from which the animal was selected*

Code A. 16S rRNA gene sequencing statistical analysis.

```
library(afex)
library(lme4)
library(emmeans)
library(lubridate)
library(ggplot2)
library("cowplot")
theme_set(theme_grey())
#install.packages("jtools")
library(jtools)
library(ggpubr)
library(sjstats)
rm(list = ls ())
```

```
setwd("~/Desktop/eunice/Thesis/Scripts/PaperScript/Metadata/")
metadata <- read.csv("16Smetadata.csv", na.strings = c("","NA"), header=TRUE)
#assign numerical values to factors
metadata$BRD <- as.factor(metadata$BRD)
levels(metadata$BRD) <- list("Healthy"="Healthy", "BRD"="BRD")
metadata$PenCode <- as.factor(metadata$PenCode)
str(metadata)

# Converting date of collection to numeric values
metadata$date <- metadata$Date.Collection
metadata$Date.Collection <- as.Date(metadata$Date.Collection, "%m/%d/%y")
d<- as.Date('12/31/2020', "%m/%d/%y") #use to calculate the days
metadata$Date.Collection <- as.Date(d) -as.Date(metadata$Date.Collection)
metadata$Date.Collection <- as.numeric(metadata$Date.Collection)
str(metadata$Date.Collection)
#the highest day value is the date of the samples collected first
plot(metadata$Date.Collection, metadata$Age)
set_sum_contrasts() # important for afex
# full model
str(metadata)

#For dependent variable Observed OTUs
#install.packages("piecewiseSEM")
M1 <- mixed(observed_otus ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method
= "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1)
M2 <- mixed(pielou_e ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2)
M3 <- mixed(chao1 ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M3)
M4 <- mixed(faith_pd ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4)

#checking assumptions
# way 1:
plot(M1$full_model)
plot(M2$full_model)
plot(M3$full_model)
```

```
plot(M4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(M1$full_model))
qqnorm(residuals(M2$full_model))
qqnorm(residuals(M3$full_model))
qqnorm(residuals(M4$full_model))

# interpreting results
# BRD plots
a <- afex_plot(M1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 4), mapping="color") +theme_bw() + theme(legend.position="bottom") + labs(y = "Observed ASVs", x = "Health Status") + theme(legend.position="none")
c <- afex_plot(M2, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 4), mapping="color") +theme_bw() + theme(legend.position="bottom") +labs(y = "Evenness (Pielou)", x = "Health Status") +theme(legend.position="none")
b <- afex_plot(M3, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 4), mapping="color") +theme_bw() + theme(legend.position="bottom") +labs(y = "Chao 1", x = "Health Status") +theme(legend.position="none")
d <- afex_plot(M4, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 4), mapping="color") +theme_bw() + theme(legend.position="bottom") +labs(y = "Faith_pd", x = "Health Status") +theme(legend.position="none")
plot_grid(a, b, c , d, labels = "AUTO")
### Date collection plots
e <- ggplot(data = metadata, aes(x = Date.Collection, y = observed_otus)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Observed ASVs") +xlab ("Days relative to the end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
f <- ggplot(data = metadata, aes(x = Date.Collection, y = chao1)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Chao1") +xlab ("Days relative to the end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
g <- ggplot(data = metadata, aes(x = Date.Collection, y = pielou_e)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Evenness (Pielou)") +xlab ("Days relative to the end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
```

```
h <- ggplot(data = metadata, aes(x = Date.Collection, y = faith_pd)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Faith_pd") +xlab ("Days relative to the end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
ggarrange(e,f,g,h, labels = c("A", "B", "C", "D"),ncol = 2, nrow=2, font.label = list(family = "Times New Roman"))

## Correlation between average temperature and date of collection
str(metadata)
cordata = metadata[,c(3,5)]
corr <- round(cor(cordata), 1)
corr
str(cordata)
cor(cordata$Date.Collection, cordata$Ave.temp)
cor.test(cordata$Date.Collection, cordata$Ave.temp)
library("ggpubr")
ggscatter(cordata, x = "Date.Collection", y = "Ave.temp", add = "reg.line", conf.int = TRUE, cor.coef = TRUE, cor.method = "pearson",xlab = "Days relative to the end of the study", ylab = "Average temperature (°F)")
## testing temperature
str(meta)
set_sum_contrasts()
M1 <- mixed(observed_otus ~ Ave.temp + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1)
M2 <- mixed(pielou_e ~ Ave.temp + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2)
M3 <- mixed(chao1 ~ Ave.temp + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M3)
M4 <- mixed(faith_pd ~ Ave.temp + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4)

##Checking assumptions
#checking assumptions
# way 1:
plot(M1$full_model)
plot(M2$full_model)
```

```
plot(M3$full_model)
plot(M4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(M1$full_model))
qqnorm(residuals(M2$full_model))
qqnorm(residuals(M3$full_model))
qqnorm(residuals(M4$full_model))

# Plots
i <- ggplot(data = metadata, aes(x = Ave.temp, y = observed_otus)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Observed ASVs") +xlab ("Average daily temperature (°F)") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
j <- ggplot(data = metadata, aes(x = Ave.temp, y = chao1)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Chao1") +xlab ("Average daily temperature (°F)") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
k <- ggplot(data = metadata, aes(x = Ave.temp, y = pielou_e)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Evenness (Pielou)") +xlab ("Average daily temperature (°F)") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
l <- ggplot(data = metadata, aes(x = Ave.temp, y = faith_pd)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() + ylab("Faith") +xlab ("Average daily temperature (°F)") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
ggarrange(i,j,k,l, labels = c("A", "B", "C", "D"),ncol = 2, nrow=2, font.label = list(family = "Times New Roman"))
```

Code B. qPCR statistical analysis

```
library(afex)
library(lme4)
library(emmeans)
library(lubridate)
library(psych)
library(sjstats)
library(tidyverse)
```

```
library(ggfortify)
library(pwr)
library(ggpubr)
# clear memory
#rm(list = ls ())
# load data

setwd("~/Desktop/eunice/Thesis/Scripts/PaperScript/Metadata/")
metadata <- read.csv("qPCRMetadata.csv", na.strings = c("","NA"), header=TRUE)
#assign numerical values to factors
metadata$BRD <- as.factor(metadata$BRD)
levels(metadata$BRD) <- list("Healthy"="Healthy", "BRD"="BRD")
metadata$PenCode <- as.factor(metadata$PenCode)
metadata$MbPre_Abs <- as.factor(metadata$MbPre_Abs)
metadata$MhPres_Abs <- as.factor(metadata$MhPres_Abs)
metadata$HsPres_Abs <- as.factor(metadata$HsPres_Abs)
metadata$PmPres_Abs <- as.factor(metadata$PmPres_Abs)
# Converting date to numeric value
metadata$date <- metadata$Date.Collection
metadata$Date.Collection <- as.Date(metadata$Date.Collection, "%m/%d/%y")
d<- as.Date('12/31/2020', "%m/%d/%y") #use to calculate the days
metadata$Date.Collection <- as.Date(d) -as.Date(metadata$Date.Collection)
metadata$Date.Collection <- as.numeric(metadata$Date.Collection)
str(metadata$Date.Collection)
#the highest day value is the date of the samples collected first
str(metadata)
#levels(metadata$Date)
plot(metadata$Date.Collection, metadata$Age)
set_sum_contrasts() # important for afex

## 16S data
str(metadata)
S1 <- mixed(X16S_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(S1)
#Checking assumptios
plot(S1$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(S1$full_model))
##Transform data
metadata <- mutate(metadata, X16S_log = log10(X16S_copies + 1))
metadata <- mutate(metadata, X16S_sqrt = sqrt(X16S_copies + 0.5))
```

```
S1.1 <- mixed(X16S_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(S1.1)
S1.2 <- mixed(X16S_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(S1.2)

#checking
plot(S1$full_model)
plot(S1.1$full_model)
plot(S1.2$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(S1$full_model))
qqnorm(residuals(S1.1$full_model)) #better
qqnorm(residuals(S1.2$full_model))
str(metadata)
afex_plot(S1.1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 4), mapping="color") +theme_bw() + theme(legend.position="bottom") +labs(y = "16S rRNA gene copy (log10)", x = "Health Status") +theme(legend.position="none")

#For dependent variable ---------qPCR
M1 <- mixed(Mbovis_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1)
M2 <- mixed(Pm_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2)
M3 <- mixed(Hs_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M3)
M4 <- mixed(Mh_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4)

#checking assumptions
# way 1:
plot(M1$full_model)
plot(M2$full_model)
plot(M3$full_model)
plot(M4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(M1$full_model)) #weird
```

```
qqnorm(residuals(M2$full_model)) #weird
qqnorm(residuals(M3$full_model))
qqnorm(residuals(M4$full_model)) #weird
##transformation
metadata <- mutate(metadata, Mbovis_log = log10(Mbovis_copies + 1))
metadata <- mutate(metadata, Mbovis_sqrt = sqrt(Mbovis_copies + 0.5))
metadata <- mutate(metadata, Pm_log = log10(Pm_copies + 1))
metadata <- mutate(metadata, Pm_sqrt = sqrt(Pm_copies + 0.5))
metadata <- mutate(metadata, Mh_log = log10(Mh_copies + 1))
metadata <- mutate(metadata, Mh_sqrt = sqrt(Mh_copies + 0.5))

## Test new variables
#For dependent variable
M1.1 <- mixed(Mbovis_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method
= "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1.1)
M1.2 <- mixed(Mbovis_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method
= "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1.2)
M2.1 <- mixed(Pm_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2.1)
M2.2 <- mixed(Pm_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2.2)
M4.1 <- mixed(Mh_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4.1)
M4.2 <- mixed(Mh_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4.2)

#checking assumptions
plot(M1$full_model)
plot(M1.1$full_model) #better
plot(M1.2$full_model)
plot(M2$full_model)
plot(M2.1$full_model)
plot(M2.2$full_model) #better
plot(M4$full_model)
plot(M4.1$full_model) #better
plot(M4.2$full_model)
```

```
# this is for testing the normality of the residuals
qqnorm(residuals(M1$full_model))
qqnorm(residuals(M1.1$full_model)) #better log
qqnorm(residuals(M1.2$full_model))
qqnorm(residuals(M2$full_model))
qqnorm(residuals(M2.1$full_model)) #better log
qqnorm(residuals(M2.2$full_model))
qqnorm(residuals(M4$full_model))
qqnorm(residuals(M4.1$full_model)) #better log
qqnorm(residuals(M4.2$full_model))
## new variables, Mbovis_log, Pm_sqrt, and Mh_log
M1.1 <- mixed(Mbovis_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method
= "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1.1)
M2.1 <- mixed(Pm_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2.1)
M3 <- mixed(Hs_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M3)
M4.1 <- mixed(Mh_log ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M4.1)

# interpreting results
afex_plot(M1.1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank())
+ labs(y = "M. bovis copies (log10)- All Samples", x = "Health Status")
+theme(legend.position="none")
afex_plot(M3, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank())
+ labs(y = "H. somni copies- All Samples", x = "Health Status") +theme(legend.position="none")
afex_plot(M4.1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank())
+ labs(y = "M. haemolytica copies (log10)-All Samples", x = "Health Status") +
theme(legend.position="none")
## Relative abundance values
m1 <- mixed(Mbovis_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data =
metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(m1)
```

```
m2 <- mixed(Pm_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(m2)
m3 <- mixed(Hs_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(m3)
m4 <- mixed(Mh_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(m4)
###Plots
afex_plot(m1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color") + theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank()) +  labs(y = "M. bovis- Rel.Abun (All Samples)", x = "Health Status") +theme(legend.position="none")
afex_plot(m2, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color") +theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank()) + labs(y = "P. multocida- Rel.Abun (All Samples)", x = "Health Status") +theme(legend.position="none")
#Date plots
a <- ggplot(data = metadata, aes(x = Date.Collection, y = Mbovis_RelAbunBacteria)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() +  ylab("M. bovis- Rel.Abun (All Samples)") +xlab ("Days relative to end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
b <- ggplot(data = metadata, aes(x = Date.Collection, y = Pm_RelAbunBacteria)) + geom_point() + geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic() +  ylab("P. multocida- Rel.Abun (All Samples)") +xlab ("Days relative to end of the study") +theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y = element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 14), axis.text.y = element_text(color = "black", size = 14))
ggarrange(a,b, labels = c("A", "B", "C", "D"), ncol = 2)

#checking assumptions
plot(m1$full_model)
plot(m2$full_model)
plot(m3$full_model)
plot(m4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(m1$full_model)) #weird
qqnorm(residuals(m2$full_model)) #weird
qqnorm(residuals(m3$full_model))
```

```
qqnorm(residuals(m4$full_model)) #weird

## Only animals that tested positive
MbYes <- subset(metadata, subset=MbPre_Abs %in% c("1"))
HsYes <- subset(metadata, subset=HsPres_Abs %in% c("1"))
MhYes <- subset(metadata, subset=MhPres_Abs %in% c("1"))
PmYes <- subset(metadata, subset=PmPres_Abs %in% c("1"))
#Copy number
y1 <- mixed(Mbovis_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = MbYes,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y1)
y2 <- mixed(Pm_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = PmYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y2)
y3 <- mixed(Hs_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = HsYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y3)
y4 <- mixed(Mh_copies ~ BRD + Date.Collection + Age + (1|PenCode), data = MhYes,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y4)

#checking assumptions
plot(y1$full_model)
plot(y2$full_model)
plot(y3$full_model)
plot(y4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(y1$full_model)) #weird
qqnorm(residuals(y2$full_model)) #weird
qqnorm(residuals(y3$full_model))
qqnorm(residuals(y4$full_model)) #weird
str(metadata)
y1.1 <- mixed(Mbovis_log ~ BRD + Date.Collection + Age + (1|PenCode), data = MbYes,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y1.1)
y1.2 <- mixed(Mbovis_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = MbYes,method =
"KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y1.2)
y2.1 <- mixed(Pm_log ~ BRD + Date.Collection + Age + (1|PenCode), data = PmYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y2.1)
```

```
y2.2 <- mixed(Pm_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data = PmYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y2.2)
y4.1 <- mixed(Mh_log ~ BRD + Date.Collection + Age + (1|PenCode), data =MhYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y4.1)
y4.2 <- mixed(Mh_sqrt ~ BRD + Date.Collection + Age + (1|PenCode), data =MhYes,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(y4.2)

#checking assumptions
plot(y1$full_model)
plot(y1.1$full_model) #better
plot(y1.2$full_model)
plot(y2$full_model)
plot(y2.1$full_model)
plot(y2.2$full_model) #better
plot(y4$full_model)
plot(y4.1$full_model) #better
plot(y4.2$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(y1$full_model))
qqnorm(residuals(y1.1$full_model)) #better log
qqnorm(residuals(y1.2$full_model))
qqnorm(residuals(y2$full_model))
qqnorm(residuals(y2.1$full_model)) #better log
qqnorm(residuals(y2.2$full_model))
qqnorm(residuals(y4$full_model))
qqnorm(residuals(y4.1$full_model)) #better log
qqnorm(residuals(y4.2$full_model))
##log values are better

# interpreting results
afex_plot(y1.1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+ theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x =
element_blank()) + labs(y = "M. bovis-(log10)- Positive", x = "Health Status") +
theme(legend.position="none")
afex_plot(y3, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+ theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x =
element_blank()) + labs(y = "H. somni copies - Positive", x = "Health Status") +
theme(legend.position="none")
```

```
afex_plot(y4.1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+ theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x =
element_blank()) +labs(y = "M. haemolytica-(log10)-Positive", x = "Health Status")
+theme(legend.position="none")
c <- ggplot(data = MbYes, aes(x = Date.Collection, y = Mbovis_log)) + geom_point() +
geom_smooth(method = "lm", se=TRUE) + scale_color_brewer(palette = "Dark2") + theme_classic()
+  ylab("M. bovis-Rel.Abund (log10)-Positive") +xlab ("Days relative to end of the study")
+theme(axis.title.x = element_text(color="black", size=14, face="bold"), axis.title.y =
element_text(color="black", size=14, face="bold")) + theme(axis.text.x = element_text(color =
"black", size = 14), axis.text.y = element_text(color = "black", size = 14))
ggarrange(a,c,b, labels = c("A", "B", "C"),ncol = 3)

## Relative abundance
Y1 <- mixed(Mbovis_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data =
MbYes,method = "KR",  control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(Y1)
Y2 <- mixed(Pm_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data =
PmYes,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(Y2)
Y3 <- mixed(Hs_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data =
HsYes,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(Y3)
str(MhYes)
Y4 <- mixed(Mh_RelAbunBacteria ~ BRD + Date.Collection + Age + (1|PenCode), data =
MhYes,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(Y4)
#Plots
afex_plot(Y1, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank())
+ labs(y = "M. bovis. Rel.Abund-(log10)- Positive", x = "Health Status")
+theme(legend.position="none")
afex_plot(Y2, x = "BRD", id = "PenCode", dodge = 0.4, point_arg = list(size = 6), mapping="color")
+theme_bw(base_size = 15) + theme(legend.position="bottom", panel.grid.major.x = element_blank())
+ labs(y = "P. multocida. Rel.Abund-(log10)- Positive", x = "Health Status")
+theme(legend.position="none")

### Average temperature and relative abundance
str(metadata)
set_sum_contrasts()
M1 <- mixed(Mbovis_RelAbunBacteria ~ Ave.temp + (1|PenCode), data = metadata,method = "KR",
control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M1)
```

```
M2 <- mixed(Pm_RelAbunBacteria ~ Ave.temp + (1|PenCode), data = metadata,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M2)
M3 <- mixed(Mbovis_RelAbunBacteria ~ Ave.temp + (1|PenCode), data = MbYes,method = "KR", control = lmerControl(optCtrl = list(maxfun = 1e6)), expand_re = TRUE)
anova(M3)

##Checking assumptions
#checking assumptions
plot(M1$full_model)
plot(M2$full_model)
plot(M3$full_model)
plot(M4$full_model)
# this is for testing the normality of the residuals
qqnorm(residuals(M1$full_model))
qqnorm(residuals(M2$full_model))
qqnorm(residuals(M3$full_model))
qqnorm(residuals(M4$full_model))
```

Code C. Co-occurrence analysis

```
library(Hmisc)
library(plyr)
library(fdrtool)
library(intergraph)
library(ggplot2)
library(tidyr)
library(lubridate)
library(cooccur)
library(qiime2R)
library(phyloseq)
library(naniar) ##for replace_with_na_all function
# this is the data

setwd("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/network/")
#OTU table (shared file)
#The OTU table as exported from qiime has a pound sign before the header row. You need to delete that pound sign in a text editor.
metadata <- read.delim("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/DESeq/DESeqmetadata.txt", sep = "\t", header = T, quote = "", stringsAsFactors = F)
#metadata <- metadata2[-1,]
str(metadata)
```

```
metadata$BRD <- factor(metadata$BRD)
order_groups <- metadata$ID
row.names(metadata) = metadata[,1]
metadata = metadata[,-1]

ASVs <- read_qza("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/Qiime/table-filtered2.qza")
ASV_s <- as.data.frame(ASVs$data)
ASV_table <- as.data.frame(ASVs$data) #18010 ASVs
ASV_table$ASVnos <- paste0("ASV", 1:nrow(ASV_table))
ASV_table$ASVstring <- rownames(ASV_table)
rownames(ASV_table) <- ASV_table$ASVnos ##We change the ASV name created in Qiime to ASVn
ASVkey <- ASV_table[, (ncol(ASV_table)-1):ncol(ASV_table)] #the key withe the names
ASV_table <- ASV_table[,-(ncol(ASV_table)-1):-ncol(ASV_table)]
ASV_table <- t(ASV_table)

#Taxonomy of each OTU
tax <- read_qza("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/Qiime/taxonomy.qza")
tax <- as.data.frame(tax$data)
tax2 = separate(tax, Taxon, into = c("Domain", "Phylum", "Class", "Order", "Family", "Genus", "Species"), sep=";")
#All the strings that need to be removed and replaced with NA
na_strings <- c(" s__", " g__", " f__", " o__", " c__")
tax3 = replace_with_na_all(tax2, condition = ~.x %in% na_strings)
#Next, all these `NA` classifications with the last level that was classified
tax3[] <- t(apply(tax3, 1, zoo::na.locf))
tax3 <- as.data.frame(tax3)
row.names(tax3) <- tax3[,1]
tax3 = tax3[,-c(1:2)]
tax.clean <- as.data.frame(tax3)
tax.clean$OTUs <- rownames(tax.clean)
###Remove all the OTUs that don't occur in our OTU.clean data set
tax.final = tax.clean[row.names(tax.clean) %in% row.names(ASV_s),]
##Remove unnecessary information from the taxonomy names
tax.final$Phylum <- sub("D_0__*", "", tax.final[,1])
tax.final$Phylum <- sub("D_1__*", "", tax.final[,1])
tax.final$Class <- sub("D_0__*", "", tax.final[,2])
tax.final$Class <- sub("D_1__*", "", tax.final[,2])
tax.final$Class <- sub("D_2__*", "", tax.final[,2])
tax.final$Order <- sub("D_0__*", "", tax.final[,3])
```

```
tax.final$Order <- sub("D_1__*", "", tax.final[,3])
tax.final$Order <- sub("D_2__*", "", tax.final[,3])
tax.final$Order <- sub("D_3__*", "", tax.final[,3])
tax.final$Family <- sub("D_0__*", "", tax.final[,4])
tax.final$Family <- sub("D_1__*", "", tax.final[,4])
tax.final$Family <- sub("D_2__*", "", tax.final[,4])
tax.final$Family <- sub("D_3__*", "", tax.final[,4])
tax.final$Family <- sub("D_4__*", "", tax.final[,4])
tax.final$Genus <- sub("D_0__*", "", tax.final[,5])
tax.final$Genus <- sub("D_1__*", "", tax.final[,5])
tax.final$Genus <- sub("D_2__*", "", tax.final[,5])
tax.final$Genus <- sub("D_3__*", "", tax.final[,5])
tax.final$Genus <- sub("D_4__*", "", tax.final[,5])
tax.final$Genus <- sub("D_5__*", "", tax.final[,5])
tax.final$Species <- sub("D_0__*", "", tax.final[,6])
tax.final$Species <- sub("D_1__*", "", tax.final[,6])
tax.final$Species <- sub("D_2__*", "", tax.final[,6])
tax.final$Species <- sub("D_3__*", "", tax.final[,6])
tax.final$Species <- sub("D_4__*", "", tax.final[,6])
tax.final$Species <- sub("D_5__*", "", tax.final[,6])
tax.final$Species <- sub("D_6__*", "", tax.final[,6])
TaxASV <- merge(tax.final, ASVkey, by.x = 0, by.y = "ASVstring")
row.names(TaxASV) <- TaxASV[,10]
TaxASV = TaxASV[,-c(1,10)]

### Creating the Phyloseq Object
OTU.physeq = otu_table(as.matrix(ASV_table), taxa_are_rows=FALSE)
tax.physeq = tax_table(as.matrix(TaxASV))
#meta.physeq = sample_data(meta)
meta.physeq = sample_data(metadata)

#We then merge these into an object of class phyloseq.
physeq_deseq = phyloseq(OTU.physeq, tax.physeq, meta.physeq)
physeq_deseq
colnames(tax_table(physeq_deseq))
## Filter any non-baxteria, chloroplast and mitochondria
physeq_deseq %>%subset_taxa(Family != "Mitochondria" & Genus != "Mitochondria" &Species !=
"Mitochondria" &Order != "Chloroplast" &Family != "Chloroplast" &Genus != "Chloroplast"
&Species != "Chloroplast") -> physeq_deseq
physeq_deseq

# Step 1: subset the samples based on healthy or sick
```

```
#Pruning the data
# Set prunescale
prunescale = 0.0001
minlib = 40420 # rarefying to 1500 reads
# Prune out rare OTUs by mean relative abundance set by prunescale
tax.mean <- taxa_sums(physeq_deseq)/nsamples(physeq_deseq)
sites.prune <- prune_taxa(tax.mean > prunescale*minlib, physeq_deseq)
sites.prune
prunetable4<- phyloseq_to_df(sites.prune, addtax = T, addtot = F, addmaxrank = F, sorting = "abundance")
#Subseting the OTU table based on healthy and sick
OTUtable <- prunetable4[,-c(2:9)]
row.names(OTUtable) <- OTUtable[,1]
OTUtable<- OTUtable[,-c(1)]
OTUtable <- t(OTUtable)
# Taxonomy table
TaxaPrune<- prunetable4[,-c(10:140)]
## Merging metadata and OTU table to then divide the data
str(metadata)
data <- merge(metadata, OTUtable, by.x = 0, by.y = 0)
str(data)

## Subset the data into healthy and sick
dataHealthy <- subset(data, BRD=="0")
dataHealthy<- dataHealthy[,-c(2:3)] #to remove any unnecessary columns
row.names(dataHealthy) <-dataHealthy[,1]
dataHealthy<- dataHealthy[,-c(1)]
dataHealthy <- t(dataHealthy)
#write.table(dataHealthy,"dataHealthy.txt",sep=",", row.names = TRUE)
dataHealthy[dataHealthy > 1] <- 1 ## we transpose because for the co-occur function, it needs the rows be ASVs and colums sites (ID)
#write.table(dataHealthy,"dataHealthyPresence.txt",sep=",", row.names = TRUE)
dataBRD <- subset(data, BRD=="1")
dataBRD<- dataBRD[,-c(2:3)] #to remove any unnecessary columns
row.names(dataBRD) <-dataBRD[,1]
dataBRD<- dataBRD[,-c(1)]
dataBRD<- t(dataBRD)
#write.table(dataBRD,"dataBRD.txt",sep=",", row.names = TRUE)

dataBRD[dataBRD > 1] <- 1## we transpose because for the co-occur function, it needs the rows be ASVs and colums sites (ID)
#write.table(dataBRD,"dataBRDPresence.txt",sep=",", row.names = TRUE)
```

```
### The co-occurrence step was run using the Purdue cluster
## Running co-occur
B <- cooccur(mat=dataBRD, type="spp_site", thresh=TRUE, spp_names=TRUE)
summary(H1co)
plot(H1co)
print(B)
prob.table(B)
H <- cooccur(mat=dataHealthy, type="spp_site", thresh=TRUE, spp_names=TRUE)
summary(H1co)
plot(H1co)
print(H)
prob.table(H)

##Cooocurrence data Healthy
CooccurH <- read.csv("printHealthy.txt", na.strings = c("","NA"), header=TRUE)
CoocpronH <- read.csv("prob.tableHealthy.txt", na.strings = c("","NA"), header=TRUE)
### CooccurH contains only the positive and negative pairs
# p_lt =1 positive
# p_gt= negative
positivepairs <-subset(CooccurH, p_lt > 0.9)
positivepairs <- merge(TaxaPrune, positivepairs, by.x = "OTU", by.y = "sp1_name")
colnames(positivepairs) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
positivepairs <- merge(TaxaPrune,positivepairs, by.x = "OTU", by.y = "sp2_name")
colnames(positivepairs) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
#write.table(positivepairs,"positivepairsHealthy.txt",sep=",", row.names = FALSE)
positiveH<-subset(CooccurH, p_lt > 0.999999) ##only the ones with high positive probability
positiveH <- subset(CooccurH, obs_cooccur > 45) ### probability that the species will be in at least 60% of the total samples
positiveH <- subset(CooccurH, prob_cooccur > 0.9) ### probability that the two species will be in the same site
str(positiveH)
PosTax <- merge(TaxaPrune, positiveH, by.x = "OTU", by.y = "sp1_name")
colnames(PosTax) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
PosTax <- merge(TaxaPrune,PosTax, by.x = "OTU", by.y = "sp2_name")
```

```
colnames(PosTax) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
#write.table(PosTax,"PosTax.txt",sep=",", row.names = FALSE)
negativepairs <-subset(CooccurH, p_lt < 0.01)
negativepairs <-subset(negativepairs, obs_cooccur <1)
negativepairs <-subset(negativepairs, prob_cooccur <0.05)
negativepairs <- merge(TaxaPrune, negativepairs, by.x = "OTU", by.y = "sp1_name")
colnames(negativepairs) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
negativepairs <- merge(TaxaPrune,negativepairs, by.x = "OTU", by.y = "sp2_name")
colnames(negativepairs) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
write.table(negativepairs,"negativepairsHealthy.txt",sep=",", row.names = FALSE)
negaH<-subset(CooccurH, p_lt < 0.05) ##only the ones with high positive probability
negaH <- merge(TaxaPrune, negaH, by.x = "OTU", by.y = "sp1_name")
colnames(negaH) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
negaH <- merge(TaxaPrune,negaH, by.x = "OTU", by.y = "sp2_name")
colnames(negaH) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
write.table(negaH,"negativeHealthy.txt",sep=",", row.names = FALSE)

##Cooocurrence data BRD
CooccurB <- read.csv("printBRD.txt", na.strings = c("","NA"), header=TRUE)
CoocpronB <- read.csv("prob.tableBRD.txt", na.strings = c("","NA"), header=TRUE)
### CooccurH contains only the positive and negative pairs
# p_lt =1 positive
# p_gt= negative
positivepairsB <-subset(CooccurB, p_lt > 0.9)
positivepairsB <- merge(TaxaPrune, positivepairsB, by.x = "OTU", by.y = "sp1_name")
colnames(positivepairsB) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
positivepairsB <- merge(TaxaPrune,positivepairsB, by.x = "OTU", by.y = "sp2_name")
```

```
colnames(positivepairsB) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
#write.table(positivepairsB,"positivepairsB.txt",sep=",", row.names = TRUE)
positiveB<-subset(CooccurB, p_lt > 0.999999) ##only the ones with high positive probability
positiveB <- subset(CooccurB, obs_cooccur > 34) ### probability that the species will be in at least 60% of the total samples
positiveB <- subset(CooccurB, prob_cooccur > 0.9) ### probability that the two species will be in the same site
str(positiveB)
PosTaxB <- merge(TaxaPrune, positiveB, by.x = "OTU", by.y = "sp1_name")
colnames(PosTaxB) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
PosTaxB <- merge(TaxaPrune,PosTaxB, by.x = "OTU", by.y = "sp2_name")
colnames(PosTaxB) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
#write.table(PosTaxB,"PosTaxB.txt",sep=",", row.names = TRUE)
negativepairsB <-subset(CooccurB, p_lt < 0.01)
negativepairsB <-subset(negativepairsB, obs_cooccur <1)
negativepairsB <-subset(negativepairsB, prob_cooccur <0.05)
negativepairsB <- merge(TaxaPrune, negativepairsB, by.x = "OTU", by.y = "sp1_name")
colnames(negativepairsB) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
negativepairsB <- merge(TaxaPrune,negativepairsB, by.x = "OTU", by.y = "sp2_name")
colnames(negativepairsB) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
write.table(negativepairsB,"negativepairsBRD.txt",sep=",", row.names = FALSE)
negaB<-subset(CooccurB, p_lt < 0.05) ##only the ones with high positive probability
negaB <- merge(TaxaPrune, negaB, by.x = "OTU", by.y = "sp1_name")
colnames(negaB) <- c("sp1_name", "Phylum_1", "Class_1", "Order_1", "Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc", "obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt", "sp2_name")
negaB <- merge(TaxaPrune,negaB, by.x = "OTU", by.y = "sp2_name")
colnames(negaB) <- c("sp2_name", "Phylum_2", "Class_2", "Order_2", "Family_2", "Genus_2", "Species_2", "Confidence_2", "OTUs_2","sp1_name", "Phylum_1", "Class_1", "Order_1",
```

```
"Family_1", "Genus_1", "Species_1", "Confidence_1", "OTUs_1", "sp1", "sp2", "sp1_inc", "sp2_inc",
"obs_cooccur", "prob_cooccur", "exp_cooccur", "p_lt", "p_gt")
write.table(negaB,"negativeBRD.txt",sep=",", row.names = FALSE)

##Function
https://rdrr.io/github/vmikk/metagMisc/src/R/phyloseq_to_df.R
phyloseq_to_df <- function(physeq, addtax = T, addtot = F, addmaxrank = F, sorting = "abundance"){
  # require(phyloseq)
  ## Data validation
  if(any(addtax == TRUE || sorting == "taxonomy")){
    if(is.null(phyloseq::tax_table(physeq, errorIfNULL = F))){
      stop("Error: taxonomy table slot is empty in the input data.\n")
    }
  }
  ## Prepare data frame
  if(taxa_are_rows(physeq) == TRUE){
    res <- data.frame(OTU = phyloseq::taxa_names(physeq), phyloseq::otu_table(physeq),
stringsAsFactors = F)
  } else {
    res <- data.frame(OTU = phyloseq::taxa_names(physeq), t(phyloseq::otu_table(physeq)),
stringsAsFactors = F)
  }
  ## Check if the sample names were silently corrected in the data.frame
  if(any(!phyloseq::sample_names(physeq) %in% colnames(res)[-1])){
    if(addtax == FALSE){
      warning("Warning: Sample names were converted to the syntactically valid column names in
data.frame. See 'make.names'.\n")
    }
    if(addtax == TRUE){
      stop("Error: Sample names in 'physeq' could not be automatically converted to the syntactically
valid column names in data.frame (see 'make.names'). Consider renaming with 'sample_names'.\n")
    }
  }
  ## Add taxonomy
  if(addtax == TRUE){
    ## Extract taxonomy table
    taxx <- as.data.frame(phyloseq::tax_table(physeq), stringsAsFactors = F)
    ## Reorder taxonomy table
    taxx <- taxx[match(x = res$OTU, table = rownames(taxx)), ]
    ## Add taxonomy table to the data
    res <- cbind(res, taxx)
    ## Add max tax rank column
```

```
    if(addmaxrank == TRUE){
      ## Determine the lowest level of taxonomic classification
      res$LowestTaxRank <- get_max_taxonomic_rank(taxx, return_rank_only = TRUE)
      ## Reorder columns (OTU name - Taxonomy - Max Rank - Sample Abundance)
      res <- res[, c("OTU", phyloseq::rank_names(physeq), "LowestTaxRank",
phyloseq::sample_names(physeq))]
    } else {
      ## Reorder columns (OTU name - Taxonomy - Sample Abundance)
      res <- res[, c("OTU", phyloseq::rank_names(physeq), phyloseq::sample_names(physeq))]
    } # end of addmaxrank
  }   # end of addtax
  ## Reorder OTUs
  if(!is.null(sorting)){
    ## Sort by OTU abundance
    if(sorting == "abundance"){
      otus <- res[, which(colnames(res) %in% phyloseq::sample_names(physeq))]
      res <- res[order(rowSums(otus, na.rm = T), decreasing = T), ]
    }
    ## Sort by OTU taxonomy
    if(sorting == "taxonomy"){
      taxtbl <- as.data.frame( phyloseq::tax_table(physeq), stringsAsFactors = F )
      ## Reorder by all columns
      taxtbl <- taxtbl[do.call(order, taxtbl), ]
      # taxtbl <- data.table::setorderv(taxtbl, cols = colnames(taxtbl), na.last = T)
      res <- res[match(x = rownames(taxtbl), table = res$OTU), ]
    }
  }
  ## Add OTU total abundance
  if(addtot == TRUE){
    res$Total <- rowSums(res[, which(colnames(res) %in% phyloseq::sample_names(physeq))])
  }
  rownames(res) <- NULL
  return(res)
}
```

Code D. LDA analysis

```
library(tidyverse)
library(modelr)
library(broom)
library(ISLR)
library(ROCR)
```

```
library(MASS)
#install.packages("ISLR")
## Test the data

setwd("~/Desktop/eunice/Thesis/qPCR/LDA_with0/")
qPCR <- read.csv("LDAMetadata.csv", na.strings = c("","NA"), header=TRUE)
str(qPCR)
##Data with 0
qPCR <- mutate(qPCR, Mbovis_copies = log10(Mbovis_copies + 1))
qPCR <- mutate(qPCR, Pm_copies = log10(Pm_copies + 1))
qPCR <- mutate(qPCR, Hs_copies = log10(Hs_copies + 1))
qPCR <- mutate(qPCR, Mh_copies = log10(Mh_copies + 1))
qPCR <- mutate(qPCR, X16S_copies = log10(X16S_copies + 1))
QPCR<- qPCR

#change the columns depending on the variables you will use to construct the model
qPCR <- QPCR[c(1,7,8)]
str(qPCR)
set.seed(123)
#training_sample2 <- sample(c(TRUE, FALSE), nrow(qPCR), replace = T, prob = c(0.6,0.4))
sample <- sample.int(n = nrow(qPCR), size = floor(.60*nrow(qPCR)), replace = F)
train <- qPCR[sample, ]
train$rownames <- rownames(train)
test  <- qPCR[-sample, ]

#Apply LDA to the training set
str(train)
train2<- train[c(1:3)]
lda.BRD <- lda(BRD ~ ., train2)
lda.BRD #show results
plot(lda.BRD)

# See if the model fits the data -- we use the predict function
#we have to use the vector we created when we run the LDA
lda.BRD2 <- predict(lda.BRD, newdata=test)
#Validation, how well the model predicts the true positives and true negatives (%)
str(test)
#chao and shannon are number
test %>% tally()
test %>% count(BRD)
train2 %>% count(BRD)
Tlda <- lda.BRD2[["class"]]
```

```
Tlda <- as.data.frame(Tlda)
Tlda %>% tally()
Tlda %>% count(Tlda)
lda.cm <- table(test$BRD, lda.BRD2$class)
cm <- as.data.frame(lda.cm)
cm

#Clasification rate
LDA_model = lda.cm %>% prop.table() %>% round(3)
LDA_model ### percent of true positives and true negatives
#Misclasification rate
test %>%mutate(lda.pred = (lda.BRD2$class)) %>%summarise(lda.error = mean(BRD != lda.pred))
# Confusion matrices
#Confusion matrix
LDA_model = lda.cm
LDA_model
```

Code E. DESeq analysis

```
library("DESeq2")
library(dplyr)
library(tidyr)
library(ape)
library(ggpubr)
library(dplyr)
library(ggplot2)
library(phyloseq)
library(plotly)
library(tidyr)
library(naniar)
library(zoo)
library(lubridate)
library(qiime2R)
#OTU table (shared file)

#The OTU table as exported from qiime has a pound sign before the header row. You need to delete that pound sign in a text editor.
metadata <- read.delim("DESeqmetadata.txt", sep = "\t", header = T, quote = "", stringsAsFactors = F)
#metadata <- metadata2[-1,]
str(metadata)
metadata$BRD <- factor(metadata$BRD)
order_groups <- metadata$ID
row.names(metadata) = metadata[,1]
```

```
metadata = metadata[,-1]

ASVs <- read_qza("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/Qiime/table-filtered2.qza")
ASV_s <- as.data.frame(ASVs$data)
ASV_table <- as.data.frame(ASVs$data) #18010 ASVs
ASV_table$ASVnos <- paste0("ASV", 1:nrow(ASV_table))
ASV_table$ASVstring <- rownames(ASV_table)
rownames(ASV_table) <- ASV_table$ASVnos ##We change the ASV name created in Qiime to ASVn
ASVkey <- ASV_table[, (ncol(ASV_table)-1):ncol(ASV_table)] #the key withe the names
ASV_table <- ASV_table[,-(ncol(ASV_table)-1):-ncol(ASV_table)]
ASV_table <- t(ASV_table)
#ASV_table2 <- ASV_table * 2 + 1

#Taxonomy of each OTU
tax <- read_qza("~/Desktop/eunice/Thesis/Qiime/Samples/OnlySamples/Filtered/Qiime/taxonomy.qza")
tax <- as.data.frame(tax$data)
tax2 = separate(tax, Taxon, into = c("Domain", "Phylum", "Class", "Order", "Family", "Genus", "Species"), sep=";")
#All the strings that need to be removed and replaced with NA
na_strings <- c(" s__", " g__", " f__", " o__", " c__")
tax3 = replace_with_na_all(tax2, condition = ~.x %in% na_strings)
#This code is great because an ASV that is unclassified at a certain level are all listed as `NA`.
#Next, all these `NA` classifications with the last level that was classified
tax3[] <- t(apply(tax3, 1, zoo::na.locf))
tax3 <- as.data.frame(tax3)
row.names(tax3) <- tax3[,1]
tax3 = tax3[,-c(1:2)]
tax.clean <- as.data.frame(tax3)
tax.clean$OTUs <- rownames(tax.clean)
#Would be good to check here to make sure the order of the two data frames was the same. You should do this on your own.
###Remove all the OTUs that don't occur in our OTU.clean data set
tax.final = tax.clean[row.names(tax.clean) %in% row.names(ASV_s),]
##Remove unneccesary information from the taxonomy names
tax.final$Phylum <- sub("D_0__*", "", tax.final[,1])
tax.final$Phylum <- sub("D_1__*", "", tax.final[,1])
tax.final$Class <- sub("D_0__*", "", tax.final[,2])
tax.final$Class <- sub("D_1__*", "", tax.final[,2])
tax.final$Class <- sub("D_2__*", "", tax.final[,2])
```

```
tax.final$Order <- sub("D_0__*", "", tax.final[,3])
tax.final$Order <- sub("D_1__*", "", tax.final[,3])
tax.final$Order <- sub("D_2__*", "", tax.final[,3])
tax.final$Order <- sub("D_3__*", "", tax.final[,3])
tax.final$Family <- sub("D_0__*", "", tax.final[,4])
tax.final$Family <- sub("D_1__*", "", tax.final[,4])
tax.final$Family <- sub("D_2__*", "", tax.final[,4])
tax.final$Family <- sub("D_3__*", "", tax.final[,4])
tax.final$Family <- sub("D_4__*", "", tax.final[,4])
tax.final$Genus <- sub("D_0__*", "", tax.final[,5])
tax.final$Genus <- sub("D_1__*", "", tax.final[,5])
tax.final$Genus <- sub("D_2__*", "", tax.final[,5])
tax.final$Genus <- sub("D_3__*", "", tax.final[,5])
tax.final$Genus <- sub("D_4__*", "", tax.final[,5])
tax.final$Genus <- sub("D_5__*", "", tax.final[,5])
tax.final$Species <- sub("D_0__*", "", tax.final[,6])
tax.final$Species <- sub("D_1__*", "", tax.final[,6])
tax.final$Species <- sub("D_2__*", "", tax.final[,6])
tax.final$Species <- sub("D_3__*", "", tax.final[,6])
tax.final$Species <- sub("D_4__*", "", tax.final[,6])
tax.final$Species <- sub("D_5__*", "", tax.final[,6])
tax.final$Species <- sub("D_6__*", "", tax.final[,6])
TaxASV <- merge(tax.final, ASVkey, by.x = 0, by.y = "ASVstring")
row.names(TaxASV) <- TaxASV[,10]
TaxASV = TaxASV[,-c(1,10)]
### Creating the Phyloseq Object
OTU.physeq = otu_table(as.matrix(ASV_table), taxa_are_rows=FALSE)
tax.physeq = tax_table(as.matrix(TaxASV))
#meta.physeq = sample_data(meta)
meta.physeq = sample_data(metadata)
#We then merge these into an object of class phyloseq.
physeq_deseq = phyloseq(OTU.physeq, tax.physeq, meta.physeq)
physeq_deseq
colnames(tax_table(physeq_deseq))
## Filter any non-baxteria, chloroplast and mitochondria
physeq_deseq %>% subset_taxa(Family != "Mitochondria" & Genus != "Mitochondria" &Species !=
"Mitochondria" &Order != "Chloroplast" & Family != "Chloroplast" &Genus != "Chloroplast"
&Species != "Chloroplast") -> physeq_deseq
physeq_deseq
#You need to run the phyloseq_to_df function
prunetable<- phyloseq_to_df(physeq_deseq, addtax = T, addtot = F, addmaxrank = F,sorting =
"abundance")
```

```
## no mitochondria or chloroplast in the data
NewTax <- prunetable[,c(1:9)]
row.names(NewTax) <- NewTax[,1]
NewTax = NewTax[,-c(1)]
NewASVTable <- prunetable[,c(1,10:140)]
row.names(NewASVTable) <- NewASVTable[,1]
NewASVTable = NewASVTable[,-c(1)]
NewASVTable = t(NewASVTable)

### Checking how the data looks
## Make a plot to see the community in the two groups
# this prunes the taxa with abundance <2%
### CALCULATION OF THE ABUNDANCE OF EACH OTU
otu.summary <- prop.table(as.matrix(NewASVTable), 1)
str(otu.summary)
otu_abund <- colSums(otu.summary)
otu_abund2 <- as.data.frame(otu_abund)
otu.summary <- rbind(otu_abund, otu.summary)
otu.summary_sorted <- otu.summary[,order(otu.summary[1,], decreasing = TRUE)]
str(otu.summary_sorted)
melt_otu <- reshape2::melt(otu.summary_sorted[, c(1:17931)]) ###TOTAL NUMBER OF OTUS
colnames(melt_otu) <- c("Sample", "ASV", "Abundance")
str(melt_otu)
levels(melt_otu$Sample)
#merging the abundance of each OTU with the metadata and the taxonomy file
meta_otu <- merge(metadata, melt_otu, by.x = 0, by.y = "Sample")
meta_otu_tax <- merge(meta_otu, NewTax, by.x = "ASV", by.y = 0)
levels(meta_otu_tax$BRD)
meta_otu_tax$Row.names <- factor(meta_otu_tax$Row.names, levels = order_groups)
summary(meta_otu_tax$Row.names) ###to check that all the samples have the same number of OTUs
(6199 total, same value from the taxonomy file)
meta_otu_tax$Family <- factor(meta_otu_tax$Family)
meta_otu_tax$Status <- factor(meta_otu_tax$BRD)
levels(meta_otu_tax$Status) <- list("Healthy"="0", "BRD"="1")
meta_otu_tax$Genus <- factor(meta_otu_tax$Genus)
meta_otu_tax$Phylum <- factor(meta_otu_tax$Phylum)
meta_otu_tax$ASV <- factor(meta_otu_tax$ASV)
str(meta_otu_tax)

### Checking the abundance of the most common taxa
## The abundance at a family level
family <- meta_otu_tax %>% group_by(Family) %>% summarise(Abundance = sum(Abundance))
```

```
attach(family)
family <- family[order(-Abundance),]
## Familly abundance only taking in consideration treatment effect in both days
family.2 <- meta_otu_tax %>% group_by(Family, Status) %>% summarise(Abundance = sum(Abundance))
attach(family.2)
family.2 <- family.2[order(-Abundance),]
### Abundance at a phlyum level
phylum <- meta_otu_tax %>% group_by(Phylum) %>% summarise(Abundance = sum(Abundance))
attach(phylum)
phylum <- phylum[order(-Abundance),]
## Phylum abundance only taking in consideration treatment effect in both days
phylum.2 <- meta_otu_tax %>% group_by(Phylum, Status) %>% summarise(Abundance = sum(Abundance))
attach(phylum.2)
phylum.2 <- phylum.2[order(-Abundance),]
## Phylum abundance only taking in consideration season effect in both days
phylum13.3 <- meta_otu_13 %>% group_by(phylum, season) %>% summarise(Abundance = sum(Abundance))
attach(phylum13.3)
phylum13.3 <- phylum13.3[order(-Abundance),]
### Abundance at a genus level
genus <- meta_otu_tax %>% group_by(Genus) %>% summarise(Abundance = sum(Abundance))
attach(genus)
genus <- genus[order(-Abundance),]
## Genus abundance only taking in consideration treatment effect
genus.2 <- meta_otu_tax %>% group_by(Genus, Status) %>% summarise(Abundance = sum(Abundance))
attach(genus.2)
genus.2 <- genus.2[order(-Abundance),]

## PHYLUM LEVEL
num_genera <- 97 # we need 100 OTUs in order to get the 25 most abundant Genus
melt_otu1 <- reshape2::melt(otu.summary_sorted[, c(1:num_genera)])
colnames(melt_otu1) <- c("Sample", "OTU", "Abundance")
#Putting it all together: merge melt_otu, metadata, taxonomy tables
meta_otu1 <- merge(metadata, melt_otu1, by.x = 0, by.y = "Sample")
meta_otu_tax1 <- merge(meta_otu1, NewTax, by.x = "OTU", by.y = 0)
meta_otu_tax1$Row.names <- factor(meta_otu_tax1$Row.names, levels = order_groups)
summary(meta_otu_tax1$Row.names) ###to check that all the samples have the same number of OTUs (346 total)
meta_otu_tax1$Family <- factor(meta_otu_tax1$Family)
```

```
meta_otu_tax1$Status <- factor(meta_otu_tax1$BRD)
levels(meta_otu_tax1$Status) <- list("Healthy"="0", "BRD"="1")
meta_otu_tax1$Phylum <- factor(meta_otu_tax1$Phylum)
meta_otu_tax1$Genus <- factor(meta_otu_tax1$Genus)
meta_otu_tax1$Family <- factor(meta_otu_tax1$Family)
levels(meta_otu_tax1$Phylum)

##Whole phylum abundance
### Calculation of the Phylum relative abundance for each time and treatment
PhylumAB <- meta_otu_tax1 %>% group_by(Row.names, Status, Date.Collection, Phylum) %>% summarise(taxa.sum = sum(Abundance)) %>%group_by(Status, Date.Collection, Phylum) %>%summarise(taxa.average = mean(taxa.sum)) ### relative abundance
str(PhylumAB)
PhylumAB$Phylum <- factor(PhylumAB$Phylum)
PhylumAB$Date.Collection <- factor(PhylumAB$Date.Collection)
levels(PhylumAB$Phylum)
levels(PhylumAB$Date.Collection)
levels(PhylumAB$Date.Collection) <- list("7/14"="7/14/20","7/21"="7/21/20","8/12"="8/12/20", "8/19"="8/19/20","8/26"="8/26/20", "9/10"="9/10/20", "9/30"="9/30/20", "10/14"="10/14/20", "10/28"="10/28/20","11/4"="11/4/20", "11/11"="11/11/20", "11/18"="11/18/20", "12/2"="12/2/20")
my_colors <- c('#a6cee3','#1f78b4','#b3df8a','#33a03c','#fb9a99','#e31a1c','#fdbf6f','#ff7f00','#cab3d6','#6a3d9a','#ffff99','#b15938', "#CBD588", "#5F7FC7", "orange","#DA5734", "#508578", "#CD9BCD","#AD6F3B", "#673770","#D14385", "#653936", "#C84348", "#8569D5", "#5E738F","#D1A33D", "#8A7C64", "#599861", "black")
#Plot the graph
ggplot(PhylumAB, aes(x = Date.Collection, y = taxa.average, fill =Phylum)) + geom_bar(stat = "identity") +theme_bw()+scale_fill_manual(values = my_colors) +facet_grid(Status~.)+ guides(fill = guide_legend(reverse = F, keywidth = 1.5, keyheight = .6, ncol = 1)) +theme(legend.text=element_text(size=8)) +theme(strip.text = element_text(size = 13, face = "bold")) + theme(legend.text = element_text(size=13)) +theme(legend.title = element_text(size = 13, face= "bold")) +theme(legend.key.size = unit(8, "point")) +theme(axis.title.x = element_text(color="black", size=13, face="bold"), axis.title.y = element_text(color="black", size=13, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 13), axis.text.y = element_text(color = "black", size = 13)) +ylab(paste0("Relative Abundance Phylum (Top 8)")) + labs(x='Date of Collection')

### FAMILY LEVEL
num_genera <- 44 # s
melt_otu3 <- reshape2::melt(otu.summary_sorted[, c(1:num_genera)])
colnames(melt_otu3) <- c("Sample", "OTU", "Abundance")
#Putting it all together: merge melt_otu, metadata, taxonomy tables
meta_otu3 <- merge(metadata, melt_otu3, by.x = 0, by.y = "Sample")
```

```
meta_otu_tax3 <- merge(meta_otu3, NewTax, by.x = "OTU", by.y = 0)
str(meta_otu_tax3)
meta_otu_tax3$Row.names <- factor(meta_otu_tax3$Row.names, levels = order_groups)
summary(meta_otu_tax3$Row.names) ###to check that all the samples have the same number of OTUs (120)
meta_otu_tax3$Family <- factor(meta_otu_tax3$Family)
meta_otu_tax3$Status <- factor(meta_otu_tax3$BRD)
levels(meta_otu_tax3$Status) <- list("Healthy"="0", "BRD"="1")
meta_otu_tax3$Phylum <- factor(meta_otu_tax3$Phylum)
meta_otu_tax3$Genus <- factor(meta_otu_tax3$Genus)
meta_otu_tax3$Family <- factor(meta_otu_tax3$Family)
levels(meta_otu_tax3$Phylum)
FamilyAB <- meta_otu_tax3 %>% group_by(Row.names, Status, Date.Collection, Family) %>% summarise(taxa.sum = sum(Abundance)) %>%group_by(Status, Date.Collection, , Family) %>% summarise(taxa.average = mean(taxa.sum))
str(FamilyAB)
FamilyAB$Family <- factor(FamilyAB$Family)
FamilyAB$Date.Collection <- factor(FamilyAB$Date.Collection)
levels(FamilyAB$Family)
levels(FamilyAB$Date.Collection)
levels(FamilyAB$Date.Collection) <- list("7/14"="7/14/20","7/21"="7/21/20","8/12"="8/12/20", "8/19"="8/19/20","8/26"="8/26/20", "9/10"="9/10/20", "9/30"="9/30/20", "10/14"="10/14/20", "10/28"="10/28/20","11/4"="11/4/20", "11/11"="11/11/20", "11/18"="11/18/20", "12/2"="12/2/20")

#Plot the graph
ggplot(FamilyAB, aes(x = Date.Collection, y = taxa.average, fill =Family)) + geom_bar(stat = "identity") +theme_bw()+scale_fill_manual(values = my_colors) +facet_grid(Status~.)+ylim(c(0,1)) +guides(fill = guide_legend(reverse = F, keywidth = 1.0, ncol = 1)) +theme(legend.text=element_text(size=8)) +guides(fill = guide_legend(reverse = F, keywidth = 1.5, keyheight = .6, ncol = 1)) +theme(legend.text=element_text(size=8)) +theme(strip.text = element_text(size = 13, face = "bold")) +theme(legend.text = element_text(size=13)) +theme(legend.title = element_text(size = 13, face= "bold")) +theme(legend.key.size = unit(8, "point")) +theme(axis.title.x = element_text(color="black", size=13, face="bold"), axis.title.y = element_text(color="black", size=13, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 13), axis.text.y = element_text(color = "black", size = 13)) + ylab(paste0("Average Relative Abundance Family (Top 20)")) + labs(x='Date of Collection')

#Genus level
num_genera <- 10 # we need 100 OTUs in order to get the 25 most abundant Genus
melt_otu4 <- reshape2::melt(otu.summary_sorted[, c(1:num_genera)])
colnames(melt_otu4) <- c("Sample", "OTU", "Abundance")
#Putting it all together: merge melt_otu, metadata, taxonomy tables
```

```
meta_otu4 <- merge(metadata, melt_otu4, by.x = 0, by.y = "Sample")
meta_otu_tax4 <- merge(meta_otu4, NewTax, by.x = "OTU", by.y = 0)
str(meta_otu_tax4)
meta_otu_tax4$Row.names <- factor(meta_otu_tax4$Row.names, levels = order_groups)
summary(meta_otu_tax4$Row.names) ###to check that all the samples have the same number of OTUs (53) t)
meta_otu_tax4$Family <- factor(meta_otu_tax4$Family)
meta_otu_tax4$Status <- factor(meta_otu_tax4$BRD)
levels(meta_otu_tax4$Status) <- list("Healthy"="0", "BRD"="1")
meta_otu_tax4$Phylum <- factor(meta_otu_tax4$Phylum)
meta_otu_tax4$Genus <- factor(meta_otu_tax4$Genus)
meta_otu_tax4$Family <- factor(meta_otu_tax4$Family)
levels(meta_otu_tax3$Phylum)
##Whole family abuundance
GenusAB <- meta_otu_tax4 %>% group_by(Row.names, Status, Date.Collection, Genus) %>% summarise(taxa.sum = sum(Abundance)) %>%group_by(Status, Date.Collection, Genus) %>%summarise(taxa.average = mean(taxa.sum))
str(PhylumAB)
GenusAB$Genus <- factor(GenusAB$Genus)
GenusAB$Date.Collection <- factor(GenusAB$Date.Collection)
levels(GenusAB$Genus)
levels(GenusAB$Date.Collection)
levels(GenusAB$Date.Collection) <- list("7/14"="7/14/20","7/21"="7/21/20","8/12"="8/12/20", "8/19"="8/19/20","8/26"="8/26/20", "9/10"="9/10/20", "9/30"="9/30/20", "10/14"="10/14/20", "10/28"="10/28/20","11/4"="11/4/20", "11/11"="11/11/20", "11/18"="11/18/20", "12/2"="12/2/20")
# PLOT FOR THE FIRST 25 GENUS
ggplot(GenusAB, aes(x = Date.Collection, y = taxa.average, fill =Genus)) + geom_bar(stat = "identity") + theme_bw()+ scale_fill_manual(values = my_colors) +facet_grid(Status~.)+ylim(c(0,1)) +guides(fill = guide_legend(reverse = F, keywidth = 1.5, keyheight = .6, ncol = 1)) + guides(fill = guide_legend(reverse = F, keywidth = 1.5, keyheight = .6, ncol = 1)) +theme(legend.text=element_text(size=8)) +theme(strip.text = element_text(size = 13, face = "bold")) +theme(legend.text = element_text(size=13)) +theme(legend.title = element_text(size = 13, face= "bold")) +theme(legend.key.size = unit(8, "point")) +theme(axis.title.x = element_text(color="black", size=13, face="bold"), axis.title.y = element_text(color="black", size=13, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 13), axis.text.y = element_text(color = "black", size = 13)) +  ylab(paste0("Average Relative Abundance Genus (Top 20)")) +  labs(x='Date of Collection')

## Running DESeq
#To use DESeq, we need no zeros in our OTU table. So we will edit the table + 1
NewASVTable2 <- NewASVTable + 1
### Creating the Phyloseq Object
```

```
OTU.physeq = otu_table(as.matrix(NewASVTable2), taxa_are_rows=FALSE)
tax.physeq = tax_table(as.matrix(NewTax))
meta.physeq = sample_data(metadata)
#We then merge these into an object of class phyloseq.
physeq_deseq = phyloseq(OTU.physeq, tax.physeq, meta.physeq)
physeq_deseq
levels(metadata$BRD)
# establishing the model
diagdds = phyloseq_to_deseq2(physeq_deseq, ~ BRD)
#PenCode needs to be factor
diagdds = DESeq(diagdds, test="Wald", fitType="parametric")
head(diagdds)
resultsNames(diagdds)
my_contrast = c("BRD", "0", "1")
res = results(diagdds, contrast = my_contrast, cooksCutoff = FALSE, alpha=0.05)
summary(res)
res
res <- as.data.frame(res)
alpha = 0.05
sigtab = res[which(res$padj < alpha), ]
sigtab = cbind(as(sigtab, "data.frame"), as(tax_table(physeq_deseq)[rownames(sigtab), ], "matrix"))
sigtab
sigtab$High_low <- ifelse(sigtab$log2FoldChange < -1.00, 'High in Healthy',ifelse(sigtab$log2FoldChange > 1.00, 'High in BRD','Mid Change'))
#write.table(sigtab,"sigtab.txt",sep=",", row.names = TRUE)
#To manke the figures
DeSeq = read.table("DESeqResultsNew.txt", header=TRUE, sep="\t")
str(DeSeq)
DeSeq$High_low <- factor(DeSeq$High_low)
DeSeq$Species <- factor(DeSeq$Species)
levels(DeSeq$Species)
levels(DeSeq$Species) <- list("Bacteroides"="Bacteroides", "Bibersteinia"="Bibersteinia", "Helcococcus ovis"="Helcococcus ovis", "Moraxella (1)"="Moraxella (1)", "Moraxella (2)"="Moraxella (2)", "Moraxella boevrei DSM 14165 (1)"="Moraxella boevrei DSM 14165 (1)","Moraxellaceae (1)"="Moraxellaceae (1)","Mycoplasma (1)"="Mycoplasma (1)", "Mycoplasma (2)"="Mycoplasma (2)", "Mycoplasma alkalescens 14918"="Mycoplasma alkalescens 14918 ","Mycoplasma arginini"="Mycoplasma arginini", "Streptococcus"="Streptococcus" , "Trueperella pyogenes" ="Trueperella pyogenes","uncultured Parvimonas"="uncultured Parvimonas", "uncultured Bergeyella"="uncultured Bergeyella", "Clostridium sensu stricto 1"="Clostridium sensu stricto 1", "Hydrogenophaga"="Hydrogenophaga", "Luteimonas"="Luteimonas", "Moraxella boevrei DSM 14165 (2)"="Moraxella boevrei DSM 14165 (2)", "Moraxellaceae (2)"="Moraxellaceae (2)",
```

```
"Mycoplasma bovirhinis"="Mycoplasma bovirhinis", "Salinicoccus"="Salinicoccus", "uncultured Gemmobacter"="uncultured Gemmobacter")
ggplot(data = DeSeq,aes(x = Species, y = log2FoldChange, group = factor(High_low))) + coord_flip() +geom_bar(stat = "identity", aes(fill = factor(High_low)), position = position_dodge(width = 0.9)) +labs(fill= "Diagnosis") +theme_bw()+ylab("Log2 Fold Change") +xlab ("Differentially ASVs") +theme(strip.text = element_text(size = 9, face = "bold")) +theme(legend.text = element_text(size=12)) +theme(legend.title = element_text(size = 12, face= "bold")) + theme(legend.key.size = unit(8, "point")) + theme(axis.title.x = element_text(color="black", size=12, face="bold"), axis.title.y = element_text(color="black", size=12, face="bold")) + theme(axis.text.x = element_text(color = "black", size = 12), axis.text.y = element_text(color = "black", size = 12))

#Functions
phyloseq_to_df <- function(physeq, addtax = T, addtot = F, addmaxrank = F, sorting = "abundance"){
  # require(phyloseq)
  ## Data validation
  if(any(addtax == TRUE || sorting == "taxonomy")){
    if(is.null(phyloseq::tax_table(physeq, errorIfNULL = F))){
      stop("Error: taxonomy table slot is empty in the input data.\n")
    }
  }
  ## Prepare data frame
  if(taxa_are_rows(physeq) == TRUE){
    res <- data.frame(OTU = phyloseq::taxa_names(physeq), phyloseq::otu_table(physeq), stringsAsFactors = F)
  } else {
    res <- data.frame(OTU = phyloseq::taxa_names(physeq), t(phyloseq::otu_table(physeq)), stringsAsFactors = F)
  }
  ## Check if the sample names were silently corrected in the data.frame
  if(any(!phyloseq::sample_names(physeq) %in% colnames(res)[-1])){
    if(addtax == FALSE){
      warning("Warning: Sample names were converted to the syntactically valid column names in data.frame. See 'make.names'.\n")
    }
    if(addtax == TRUE){
      stop("Error: Sample names in 'physeq' could not be automatically converted to the syntactically valid column names in data.frame (see 'make.names'). Consider renaming with 'sample_names'.\n")
    }
  }
  ## Add taxonomy
  if(addtax == TRUE){
    ## Extract taxonomy table
```

```
    taxx <- as.data.frame(phyloseq::tax_table(physeq), stringsAsFactors = F)
    ## Reorder taxonomy table
    taxx <- taxx[match(x = res$OTU, table = rownames(taxx)), ]
    ## Add taxonomy table to the data
    res <- cbind(res, taxx)
    ## Add max tax rank column
    if(addmaxrank == TRUE){
      ## Determine the lowest level of taxonomic classification
      res$LowestTaxRank <- get_max_taxonomic_rank(taxx, return_rank_only = TRUE)
      ## Reorder columns (OTU name - Taxonomy - Max Rank - Sample Abundance)
      res <- res[, c("OTU", phyloseq::rank_names(physeq), "LowestTaxRank",
phyloseq::sample_names(physeq))]
    } else {
      ## Reorder columns (OTU name - Taxonomy - Sample Abundance)
      res <- res[, c("OTU", phyloseq::rank_names(physeq), phyloseq::sample_names(physeq))]
    } # end of addmaxrank
  }   # end of addtax
  ## Reorder OTUs
  if(!is.null(sorting)){
    ## Sort by OTU abundance
    if(sorting == "abundance"){
      otus <- res[, which(colnames(res) %in% phyloseq::sample_names(physeq))]
      res <- res[order(rowSums(otus, na.rm = T), decreasing = T), ]
    }
    ## Sort by OTU taxonomy
    if(sorting == "taxonomy"){
      taxtbl <- as.data.frame( phyloseq::tax_table(physeq), stringsAsFactors = F )
      ## Reorder by all columns
      taxtbl <- taxtbl[do.call(order, taxtbl), ]
      # taxtbl <- data.table::setorderv(taxtbl, cols = colnames(taxtbl), na.last = T)
      res <- res[match(x = rownames(taxtbl), table = res$OTU), ]
    }
  }
  ## Add OTU total abundance
  if(addtot == TRUE){
    res$Total <- rowSums(res[, which(colnames(res) %in% phyloseq::sample_names(physeq))])
  }
  rownames(res) <- NULL
  return(res)
}
```

Code F. Qiime2 code to process the 16S rRNA gene sequencing

Getting set up
You will need the necessary modules to run qiime
$ module load bioinfo
$ module load mothur
$ mothur
 make.file(inputdir=., type=gz, prefix=BRD.stability)
CHANGe the first row to (use “Tab” to separate it)
sample-id forward-absolute-filepath reverse-absolute-filepath
$ module load Qiime/2-2020.2
$ module list
$ qiime tools import \
--type 'SampleData[PairedEndSequencesWithQuality]' \
--input-path BRD_onlysamples.stability.files \
--input-format PairedEndFastqManifestPhred33V2 \
--output-path ./Qiime_new/demux-paired-end.qza
$ qiime demux summarize \
  --i-data ./Qiime_new/demux-paired-end.qza \
  --o-visualization ./Qiime_new/demux-paired-end.qzv
$ qiime dada2 denoise-paired \
  --i-demultiplexed-seqs ./Qiime_new/demux-paired-end.qza \
  --p-trim-left-f 0 \
  --p-trim-left-r 0 \
  --p-trunc-len-f 251 \
  --p-trunc-len-r 223 \
  --o-table ./Qiime_new/table.qza \
  --o-representative-sequences ./Qiime_new/rep-seqs.qza \
  --o-denoising-stats ./Qiime_new/denoising-stats.qza
FeatureTable and FeatureData summaries
$ qiime feature-table summarize \
  --i-table ./Qiime_new/table.qza \
  --o-visualization ./Qiime_new/table.qzv \
  --m-sample-metadata-file BRD.samplesQiime.txt
$ qiime feature-table tabulate-seqs \
  --i-data ./Qiime_new/rep-seqs.qza \
  --o-visualization ./Qiime_new/rep-seqs.qzv
$ qiime metadata tabulate \
  --m-input-file ./Qiime_new/denoising-stats.qza \
  --o-visualization ./Qiime_new/denoising-stats.qzv
$ qiime phylogeny align-to-tree-mafft-fasttree \
  --i-sequences ./Qiime_new/rep-seqs.qza \
  --o-alignment ./Qiime_new/aligned-rep-seqs.qza \
  --o-masked-alignment ./Qiime_new/masked-aligned-rep-seqs.qza \

```
  --o-tree ./Qiime_new/unrooted-tree.qza \
  --o-rooted-tree ./Qiime_new/rooted-tree.qza
$ qiime diversity alpha-rarefaction \
  --i-table ./Qiime_new/table.qza \
  --i-phylogeny ./Qiime_new/rooted-tree.qza \
  --p-max-depth 35119 \
  --m-metadata-file BRD.samplesQiime.txt   \
  --o-visualization ./Qiime_new/alpha-rarefaction-35119.qzv
$ qiime diversity alpha-rarefaction \
  --i-table ./Qiime_new/table.qza \
  --i-phylogeny ./Qiime_new/rooted-tree.qza \
  --p-max-depth 40420 \
  --m-metadata-file BRD.samplesQiime.txt   \
  --o-visualization ./Qiime_new/alpha-rarefaction-40420.qzv
$ qiime diversity core-metrics-phylogenetic \
  --i-phylogeny ./Qiime_new/rooted-tree.qza \
  --i-table ./Qiime_new/table.qza \
  --p-sampling-depth 40420 \
  --m-metadata-file BRD.samplesQiime.txt \
  --output-dir ./Qiime_new/core-metrics-results
Taxonomic analysis
$ qiime feature-classifier classify-sklearn \
  --i-classifier silva-132-99-515-806-nb-classifier.qza \
  --i-reads ./Qiime_new/rep-seqs.qza \
  --o-classification ./Qiime_new/taxonomy.qza
$ qiime metadata tabulate \
  --m-input-file ./Qiime_new/taxonomy.qza \
  --o-visualization ./Qiime_new/taxonomy.qzv
$ qiime taxa barplot \
  --i-table ./Qiime_new/table.qza \
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --m-metadata-file BRD.samplesQiime.txt \
  --o-visualization ./Qiime_new/taxa-bar-plots.qzv
## Export the documents from the cluster to the computer
$ qiime tools export --input-path ./Qiime_new/taxonomy.qza --output-path ./Qiime_new/core-
metrics-results/exported/
```

You are in core-metric-results-2 directory

```
$ qiime tools export --input-path ./rarefied_table.qza --output-path exported/
$ qiime tools export --input-path ./faith_pd_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/faith_pd.tsv
$ qiime tools export --input-path ./shannon_vector.qza --output-path exported/
```

```
$ mv exported/alpha-diversity.tsv exported/shannon.tsv
$ qiime tools export --input-path ./observed_otus_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/observed_otus.tsv
$ ls exported
$ qiime tools export --input-path ./evenness_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/evenness.tsv
$ biom convert -i exported/feature-table.biom -o exported/rarified-table.tsv --to-tsv
```

Export beta diversity, you have to move to the OnlySamples directory

```
$ qiime tools export \
  --input-path ./Qiime_new/core-metrics-results/unweighted_unifrac_pcoa_results.qza \
  --output-path ./Qiime_new/core-metrics-results/exported-unweighted_unifrac
$ mv ./Qiime_new/core-metrics-results/exported-
unweighted_unifrac/ordination.txt ./Qiime_new/core-metrics-results/exported-
unweighted_unifrac/ordination_unweighted.txt
$ qiime tools export \
  --input-path ./Qiime_new/core-metrics-results/bray_curtis_pcoa_results.qza \
  --output-path ./Qiime_new/core-metrics-results/exported-bray_curtis
$ mv ./Qiime_new/core-metrics-results/exported-bray_curtis/ordination.txt ./Qiime_new/core-
metrics-results/exported-bray_curtis/ordination_braycurtis.txt
$ qiime tools export \
  --input-path ./Qiime_new/core-metrics-results/weighted_unifrac_pcoa_results.qza \
  --output-path ./Qiime_new/core-metrics-results/exported-weighted_unifrac
$ mv ./Qiime_new/core-metrics-results/exported-weighted_unifrac/ordination.txt ./Qiime_new/core-
metrics-results/exported-weighted_unifrac/ordination_weighted.txt
$ qiime tools export \
  --input-path ./Qiime_new/core-metrics-results/weighted_unifrac_distance_matrix.qza \
  --output-path ./Qiime_new/core-metrics-results/exported-weighted_distance
$ mv ./Qiime_new/core-metrics-results/exported-weighted_distance/distance-
matrix.tsv ./Qiime_new/core-metrics-results/exported-weighted_distance/weighted-distance-
matrix.tsv
#### Removing Pseudoalteromonas and Vibrio from the data
### Revoming the Pseudoalteromonas from the samples
$ qiime taxa filter-table \
    --i-table ./Qiime_new/table.qza \
    --i-taxonomy ./Qiime_new/taxonomy.qza \
    --p-mode exact \
    --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Alteromonadales;D_4__P
seudoalteromonadaceae;D_5__Pseudoalteromonas" \
    --o-filtered-table ./Qiime_new/table-no-pseudo.qza
#removing Vibrio
```

```
$ qiime taxa filter-table \
    --i-table ./Qiime_new/table-no-pseudo.qza \
    --i-taxonomy ./Qiime_new/taxonomy.qza \
    --p-mode exact \
    --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio" \
    --o-filtered-table ./Qiime_new/table-filtered.qza
$ qiime taxa filter-table \
    --i-table ./Qiime_new/table-filtered.qza\
    --i-taxonomy ./Qiime_new/taxonomy.qza \
    --p-mode exact \
    --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio;D_6__Vibrio sp. RCB484" \
    --o-filtered-table ./Qiime_new/table-filtered2.qza
$ qiime taxa filter-table \
    --i-table ./Qiime_new/table-filtered2.qza\
    --i-taxonomy ./Qiime_new/taxonomy.qza \
    --p-mode exact \
    --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio;D_6__uncultured Shewanella sp." \
    --o-filtered-table ./Qiime_new/table-filtered2.qza
## Creating the new table.qza file
$ qiime feature-table summarize \
  --i-table ./Qiime_new/table-filtered2.qza \
  --o-visualization ./Qiime_new/table-filtered2.qzv \
  --m-sample-metadata-file BRD.samplesQiime.txt
## Filtering the representative sequences
##You need to download first the taxonomy metadata because the ID it's going to be use for the
filtering step
## The id for Pseudoalteromonas
$ qiime taxa filter-seqs \
  --i-sequences ./Qiime_new/rep-seqs.qza \
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Alteromonadales;D_4__P
seudoalteromonadaceae;D_5__Pseudoalteromonas" \
  --o-filtered-sequences ./Qiime_new/rep-seqs-nopseudo.qza
$ qiime taxa filter-seqs \
  --i-sequences ./Qiime_new/rep-seqs-nopseudo.qza \
```

```
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio" \
  --o-filtered-sequences ./Qiime_new/rep-seqs-filtered.qza
$ qiime taxa filter-seqs \
  --i-sequences ./Qiime_new/rep-seqs-filtered.qza \
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio;D_6__Vibrio sp. RCB484" \
  --o-filtered-sequences ./Qiime_new/rep-seqs-filtered2.qza
$ qiime taxa filter-seqs \
  --i-sequences ./Qiime_new/rep-seqs-filtered.qza \
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio;D_6__uncultured Shewanella sp." \
  --o-filtered-sequences ./Qiime_new/rep-seqs-filtered2.qza
$ qiime taxa filter-seqs \
  --i-sequences ./Qiime_new/rep-seqs-filtered2.qza \
  --i-taxonomy ./Qiime_new/taxonomy.qza \
  --p-exclude
"D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Vibrionales;D_4__Vibrio
naceae;D_5__Vibrio;D_6__uncultured Shewanella sp." \
  --o-filtered-sequences ./Qiime_new/rep-seqs-filtered2.qza
$ qiime feature-table tabulate-seqs \
  --i-data ./Qiime_new/rep-seqs-filtered2.qza\
  --o-visualization ./Qiime_new/rep-seqs-filtered2.qzv
$ qiime phylogeny align-to-tree-mafft-fasttree \
  --i-sequences ./Qiime_new/rep-seqs-filtered2.qza \
  --o-alignment ./Qiime_filtered/aligned-rep-seqs.qza \
  --o-masked-alignment ./Qiime_filtered/masked-aligned-rep-seqs.qza \
  --o-tree ./Qiime_filtered/unrooted-tree.qza \
  --o-rooted-tree ./Qiime_filtered/rooted-tree.qza
```

Test the rarefaction number 36221

```
$ qiime diversity alpha-rarefaction \
  --i-table ./Qiime_new/table-filtered2.qza \
  --i-phylogeny ./Qiime_filtered/rooted-tree.qza \
  --p-max-depth 40164 \
  --m-metadata-file BRD.samplesQiime.txt \
```

```
  --o-visualization ./Qiime_filtered/alpha-rarefaction-40164.qzv

### Rarefying at 1897
$ qiime diversity core-metrics-phylogenetic \
  --i-phylogeny ./Qiime_filtered/rooted-tree.qza \
  --i-table ./Qiime_new/table-filtered2.qza \
  --p-sampling-depth 40164 \
  --m-metadata-file BRD.samplesQiime.txt  \
  --output-dir ./Qiime_filtered/core-metrics-results
```

You are in core-metric-NoPseudo2 directory

```
$ qiime tools export --input-path ./rarefied_table.qza --output-path exported/
$ qiime tools export --input-path ./faith_pd_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/faith_pd.tsv
$ qiime tools export --input-path ./shannon_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/shannon.tsv
$ ls exported
$ biom convert -i exported/feature-table.biom -o exported/rarified-table.tsv --to-tsv
$ qiime tools export --input-path observed_otus_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/observed_otus.tsv
$ qiime tools export --input-path evenness_vector.qza --output-path exported/
$ mv exported/alpha-diversity.tsv exported/evenness.tsv
$ qiime tools export --input-path ./Qiime_filtered/taxonomy.qza --output-path ./Qiime_filtered/
```

Export beta diversity, you have to move to the OnlySamples directory

```
$ qiime tools export \
  --input-path ./Qiime_filtered/core-metrics-results/unweighted_unifrac_pcoa_results.qza \
  --output-path ./Qiime_filtered/core-metrics-results/exported-unweighted_unifrac
$ mv ./Qiime_filtered/core-metrics-results/exported-
unweighted_unifrac/ordination.txt ./Qiime_filtered/core-metrics-results/exported-
unweighted_unifrac/ordination_unweighted.txt
$ qiime tools export \
  --input-path ./Qiime_filtered/core-metrics-results/bray_curtis_pcoa_results.qza \
  --output-path ./Qiime_filtered/core-metrics-results/exported-bray_curtis
$ mv ./Qiime_filtered/core-metrics-results/exported-bray_curtis/ordination.txt ./Qiime_filtered/core-
metrics-results/exported-bray_curtis/ordination_braycurtis.txt
$ qiime tools export \
  --input-path ./Qiime_filtered/core-metrics-results/weighted_unifrac_pcoa_results.qza \
  --output-path ./Qiime_filtered/core-metrics-results/exported-weighted_unifrac
$ mv ./Qiime_filtered/core-metrics-results/exported-
weighted_unifrac/ordination.txt ./Qiime_filtered/core-metrics-results/exported-
weighted_unifrac/ordination_weighted.txt
```

```
qiime tools export \
  --input-path ./Qiime_filtered/core-metrics-results/weighted_unifrac_distance_matrix.qza \
  --output-path ./Qiime_filtered/core-metrics-results/exported-weighted_distance
$ mv ./Qiime_filtered/core-metrics-results/exported-weighted_distance/distance-
matrix.tsv ./Qiime_filtered/core-metrics-results/exported-weighted_distance/weighted-distance-
matrix.tsv
$ qiime tools export \
  --input-path ./Qiime_filtered/core-metrics-results/bray_curtis_distance_matrix.qza \
  --output-path ./Qiime_filtered/core-metrics-results/exported-bray_curtis_distance
mv ./Qiime_filtered/core-metrics-results/exported-bray_curtis_distance/distance-
matrix.tsv ./Qiime_filtered/core-metrics-results/exported-bray_curtis_distance/bray_curtis-distance-
matrix.tsv
Taxonomic analysis
$ qiime feature-classifier classify-sklearn \
  --i-classifier silva-132-99-515-806-nb-classifier.qza \
  --i-reads ./Qiime_new/rep-seqs-filtered2.qza \
  --o-classification ./Qiime_filtered/taxonomy.qza
$ qiime metadata tabulate \
  --m-input-file ./Qiime_filtered/taxonomy.qza \
  --o-visualization ./Qiime_filtered/taxonomy.qzv
$ qiime taxa barplot \
  --i-table ./Qiime_new/table-filtered2.qza \
  --i-taxonomy ./Qiime_filtered/taxonomy.qza \
  --m-metadata-file BRD.samplesQiime.txt \
  --o-visualization ./Qiime_filtered/taxa-bar-plots.qzv
## Getting Chao1
qiime diversity alpha \
--i-table ./Qiime_filtered/core-metrics-results/rarefied_table.qza \
--p-metric chao1 \
--o-alpha-diversity ./Qiime_filtered/chao1.qza
```

REFERENCES

Abutarbush, S. M., Pollock, C. M., Wildman, B. K., Perrett, T., Schunicht, O. C., Kent Fenton, R., Hannon, S. J., Vogstad, A. R., Kee Jim, G., & Booker, C. W. (2012). Evaluation of the diagnostic and prognostic utility of ultrasonography at first diagnosis of presumptive bovine respiratory disease. *Canadian Journal of Veterinary Research*, *76*(1), 23–32.

Agnes, J. T., Zekarias, B., Shao, M., Anderson, M. L., Gershwin, L. J., & Corbeil, L. B. (2013). Bovine respiratory syncytial virus and Histophilus somni interaction at the alveolar barrier. *Infection and Immunity*, *81*(7), 2592–2597. https://doi.org/10.1128/IAI.00108-13

Ackermann, M. R., Derscheid, R., & Roth, J. A. (2010). Innate immunology of bovine respiratory disease. In *Veterinary Clinics of North America - Food Animal Practice*, *26*(2), 215–228. Elsevier. https://doi.org/10.1016/j.cvfa.2010.03.001

Al-Ani, Z., Suut, S., & Khan, A. N. (2015). *Multifocal Lung Consolidation: Differential diagnosis & The Role of Imaging*. *June*. https://www.researchgate.net/publication/279566912_Multifocal_Lung_Consolidation_Differential_diagnosis_The_Role_of_Imaging

Amat, S., Subramanian, S., Timsit, E., & Alexander, T. W. (2017). Probiotic bacteria inhibit the bovine respiratory pathogen *Mannheimia haemolytica* serotype 1 *in vitro*. *Letters in Applied Microbiology*, *64*(5), 343–349. https://doi.org/10.1111/lam.12723

Amat, S, Holman, D. B., Timsit, E., Schwinghamer, T., & Alexander, T. W. (2019a). Evaluation of the Nasopharyngeal Microbiota in Beef Cattle Transported to a Feedlot, With a Focus on Lactic Acid-Producing Bacteria. *Frontiers in Microbiology*, *10*, 1988. https://doi.org/10.3389/fmicb.2019.01988

Amat. S, Timsit, E., Baines, D., Yanke, J., & Alexander, T. W. (2019b). Development of bacterial therapeutics against the bovine respiratory pathogen Mannheimia haemolytica. *Applied and Environmental Microbiology*, *85*(21). https://doi.org/10.1128/AEM.01359-19

Angen, Ø., Thomsen, J., Larsen, L. E., Larsen, J., Kokotovic, B., Heegaard, P. M. H., & Enemark, J. M. D. (2009). Respiratory disease in calves: Microbiological investigations on trans-tracheally aspirated bronchoalveolar fluid and acute phase protein response. *Veterinary Microbiology*, *137*(1–2), 165–171. https://doi.org/10.1016/j.vetmic.2008.12.024

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome*

Biology, *11*(10), 1–12. https://doi.org/10.1186/gb-2010-11-10-r106

Avra, T. D., Abell, K. M., Shane, D. D., Theurer, M. E., Larson, R. L., & White, B. J. (2017). A retrospective analysis of risk factors associated with bovine respiratory disease treatment failure in feedlot cattle. *Journal of Animal Science*, *95*(4), 1521–1527. https://doi.org/10.2527/jas2016.1254

Babcock, A. H., Renter, D. G., White, B. J., Dubnicka, S. R., & Scott, H. M. (2010). Temporal distributions of respiratory disease events within cohorts of feedlot cattle and associations with cattle health and performance indices. *Preventive Veterinary Medicine*, *97*(3–4), 198–219. https://doi.org/10.1016/j.prevetmed.2010.09.003

Bell, C. J., Blackburn, P., Elliott, M., Patterson, T. I. A. P., Ellison, S., Lahuerta-Marin, A., & Ball, H. J. (2014). Investigation of polymerase chain reaction assays to improve detection of bacterial involvement in bovine respiratory disease. *Journal of Veterinary Diagnostic Investigation*, *26*(5), 631–634. https://doi.org/10.1177/1040638714540166

Beers, M. H., Berkow, R., & Merck Research Laboratories. (1999). The Merck manual of diagnosis and therapy.

Bissett, A., Brown, M. V., Siciliano, S. D., & Thrall, P. H. (2013). Microbial community responses to anthropogenically induced environmental change: Towards a systems approach. *Ecology Letters*, *16*(SUPPL.1), 128–139. https://doi.org/10.1111/ele.12109

Blakebrough-Hall, C., Dona, A., D'occhio, M. J., McMeniman, J., & González, L. A. (2020). Diagnosis of Bovine Respiratory Disease in feedlot cattle using blood 1H NMR metabolomics. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-019-56809-w

Booker, C. W., Abutarbush, S. M., Morley, P. S., Kee Jim, G., Pittman, T. J., Schunicht, O. C., Perrett, T., Wildman, B. K., Kent Fenton, R., Timothy Guichon, P., & Janzen, E. D. (2008). Microbiological and histopathological findings in cases of fatal bovine respiratory disease of feedlot cattle in Western Canada - PubMed. *Canadian Veterinary Journal*, *49*(5), 473–481. https://pubmed.ncbi.nlm.nih.gov/18512458/

Bowland, S. L., & Shewen, P. E. (2000). Bovine respiratory disease: Commercial vaccines currently available in Canada. *Canadian Veterinary Journal*, *41*(1), 33–48.

Buckham Sporer, K. R., Weber, P. S. D., Burton, J. L., Earley, B., & Crowe, M. A. (2008). Transportation of young beef bulls alters circulating physiological parameters that may be effective biomarkers of stress. *Journal of Animal Science*, *86*(6), 1325–1334.

https://doi.org/10.2527/jas.2007-0762

Call, D. R., Davis, M. A., & Sawant, A. A. (2008). Antimicrobial resistance in beef and dairy cattle production. In *Animal health research reviews / Conference of Research Workers in Animal Diseases*, *9*(2), 159–167. Cambridge University Press. https://doi.org/10.1017/S1466252308001515

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581. https://doi.org/10.1038/nmeth.3869

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., Narwani, A., MacE, G. M., Tilman, D., Wardle, D. A., Kinzig, A. P., Daily, G. C., Loreau, M., Grace, J. B., Larigauderie, A., Srivastava, D. S., & Naeem, S. (2012). Biodiversity loss and its impact on humanity. *Nature*, *486*(7401), 59–67. https://doi.org/10.1038/nature11148

Caswell, J. L. (2014). Failure of Respiratory Defenses in the Pathogenesis of Bacterial Pneumonia of Cattle. Veterinary Pathology, 51(2), 393–409. https://doi.org/10.1177/0300985813502821

Caswell, Jeff L., & Archambault, M. (1996). Mycoplasma bovis pneumonia in cattle. In *Animal Health Research Reviews*, *8*(2), 161–186. https://doi.org/10.1017/S1466252307001351

Chamorro, M. F., & Palomares, R. A. (2020). Bovine Respiratory Disease Vaccination Against Viral Pathogens Modified-Live Versus Inactivated Antigen Vaccines, Intranasal Versus Parenteral, What Is the Evidence? *Veterinary Clinics of North America - Food Animal Practice*, *36*, 461–472. https://doi.org/10.1016/j.cvfa.2020.03.006

Chao, A. (1984). Nonparametric Estimation of the Number of Classes in a Population. *Scandinavian Journal of Statistics*, *11*(4), 265–270.

Chirase, N. K., & Greene, L. W. (2000). Influence of oral natural interferon-alpha on performance and rectal temperature of newly received beef steers. In PROCEEDINGS-AMERICAN SOCIETY OF ANIMAL SCIENCE WESTERN SECTION (Vol. 51, pp. 411-414).

Corbeil, L. B., Woodward, W., Ward, A. C. S., Mickelsen, W. D., & Paisley, L. (1985). Bacterial interactions in bovine respiratory and reproductive infections. *Journal of Clinical Microbiology*, *21*(5), 803–807. https://doi.org/10.1128/jcm.21.5.803-807.1985

Cortese, V. S., Braun, D. A., Crouch, D., Townsend, C., & Zukowski, B. (2012). Case report - peracute to acute fatal pneumonia in cattle caused by Bibersteinia trehalosi. Bovine Practitioner, 46(2), 138–142.

Dabo, S. M., Taylor, J. D., & Confer, A. W. (2008). Pasteurella multocida and bovine respiratory disease. *Animal Health Research Reviews*, *8*(2), 129–150. https://doi.org/10.1017/S1466252307001399

Dassanayake, R. P., Call, D. R., Sawant, A. A., Carol Casavant, N., Weiser, G. C., Knowles, D. P., & Srikumaran, S. (2010). Bibersteinia trehalosi Inhibits the growth of mannheimia haemolytica by a proximity-dependent mechanism. *Applied and Environmental Microbiology*, *76*(4), 1008–1013. https://doi.org/10.1128/AEM.02086-09

Dedonder, K. D., & Apley, M. D. (2015). A literature review of antimicrobial resistance in Pathogens associated with bovine respiratory disease. In *Animal Health Research Reviews, 16(*2), 125–134. Cambridge University Press. https://doi.org/10.1017/S146625231500016X

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology*, *72*(7), 5069–5072. https://doi.org/10.1128/AEM.03006-05

Dickson, R. P., Erb-Downward, J. R., Prescott, H. C., Martinez, F. J., Curtis, J. L., Lama, V. N., & Huffnagle, G. B. (2014). Analysis of Culture-Dependent versus Culture-Independent Techniques for Identification of Bacteria in Clinically Obtained Bronchoalveolar Lavage Fluid Downloaded from. *Journal of Clinical Microbiology*, *52*(10), 2020. https://doi.org/10.1128/JCM.01028-14

Doak, D. F., Bigger, D., Harding, E. K., Marvier, M. A., O'Malley, R. E., & Thomson, D. (1998). The statistical inevitability of stability-diversity relationships in community ecology. *American Naturalist*, *151*(3), 264–276. https://doi.org/10.1086/286117

Doyle, D., Credille, B., Lehenbauer, T. W., Berghaus, R., Aly, S. S., Champagne, J., Blanchard, P., Crossley, B., Berghaus, L., Cochran, S., & Woolums, A. (2017). Agreement Among 4 Sampling Methods to Identify Respiratory Pathogens in Dairy Calves with Acute Bovine Respiratory Disease. *Journal of Veterinary Internal Medicine*, *31*(3), 954–959. https://doi.org/10.1111/jvim.14683

Duff, G. C., & Galyean, M. L. (2007). Board-invited review: Recent advances in management of

highly stressed, newly received feedlot cattle. *Journal of Animal Science, 85*(3), pp. 823–840). https://doi.org/10.2527/jas.2006-501

Dunlap, P. V. (2001). Microbial Diversity. In Encyclopedia of Biodiversity: Second Edition (pp. 280–291). Elsevier Inc. https://doi.org/10.1016/B978-0-12-384719-5.00435-4

Edward, A. J. (1996). Respiratory disease in feedlot cattle in the central USA. *Bovine Pract.*, *30*, 5–7. https://doi.org/10.21423/bovine-vol1996no30p5-7

Edwards, T. A. (2010). Control methods for bovine respiratory disease for feedlot cattle. *Veterinary Clinics of North America - Food Animal Practice*, *26*(2), 273–284. https://doi.org/10.1016/j.cvfa.2010.03.005

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1), 1–10. https://doi.org/10.1016/0006-3207(92)91201-3

Faner, R., Sibila, O., Agustí, A., Bernasconi, E., Chalmers, J. D., Huffnagle, G. B., Manichanh, C., Molyneaux, P. L., Paredes, R., Brocal, V. P., Ponomarenko, J., Sethi, S., Dorca, J., & Monsó, E. (2017). The microbiome in respiratory medicine: Current challenges and future perspectives. *European Respiratory Journal*, *49*(4), 1–12. https://doi.org/10.1183/13993003.02086-2016

Fulton, R. (2003). Respiratory disease in cattle: Isolation of infectious agents and lesions in fatal feedlot cases. Proceedings of the Academy of Veterinary Consultants, *31*(3), 57.

Fulton, R. W., & Confer, A. W. (2012). Laboratory test descriptions for bovine respiratory disease diagnosis and their strengths and weaknesses: Gold standards for diagnosis, do they exist? In *Canadian Veterinary Journal*, *53*(7), 754–761. Canadian Veterinary Medical Association. /pmc/articles/PMC3377458/?report=abstract

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. *Canadian Journal of Veterinary Research*, *66*(3), 173–180. /pmc/articles/PMC227001/?report=abstract

Galkiewicz, J. P., & Kellogg, C. A. (2008). Cross-kingdom amplification using Bacteria-specific primers: Complications for studies of coral microbial ecology. *Applied and Environmental Microbiology*, *74*(24), 7828–7831. https://doi.org/10.1128/AEM.01303-08

Gentry, T. J., Pepper, I. L., & Pierson, L. S. (2015). Microbial Diversity and Interactions in Natural

Ecosystems. In *Environmental Microbiology: Third Edition*. Elsevier Inc. https://doi.org/10.1016/B978-0-12-394626-3.00019-3

Goff, J. P., & Horst, R. L. (1997). Physiological Changes at Parturition and Their Relationship to Metabolic Disorders. *Journal of Dairy Science*, *80*(7), 1260–1268. https://doi.org/10.3168/jds.S0022-0302(97)76055-7

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS ONE*, *12*(5), 1–26. https://doi.org/10.1371/journal.pone.0177675

Gorden, P. J., & Plummer, P. (2010). Control, management, and prevention of bovine respiratory disease in dairy calves and cows. In *Veterinary Clinics of North America - Food Animal Practice*, *26*(2), 243–259. Elsevier. https://doi.org/10.1016/j.cvfa.2010.03.004

Griffin, D., Chengappa, M. M., Kuszak, J., & McVey, D. S. (2010). Bacterial pathogens of the bovine respiratory disease complex. In *Veterinary Clinics of North America - Food Animal Practice* (Vol. 26, Issue 2, pp. 381–394). Vet Clin North Am Food Anim Pract. https://doi.org/10.1016/j.cvfa.2010.04.004

Griffith. D, A. M., Veech, J. A., & Marsh Maintainer Daniel M Griffith, C. J. (2016). *Package "cooccur" Type Package Title Probabilistic Species Co-Occurrence Analysis in R*. https://doi.org/10.1111/j.1466-8238.2012.00789.x

Guenther, S., Schierack, P., Grobbel, M., Lübke-Becker, A., Wieler, L. H., & Ewers, C. (2008). Real-time PCR assay for the detection of species of the genus Mannheimia. *Journal of Microbiological Methods*, *75*(1), 75–80. https://doi.org/10.1016/j.mimet.2008.05.008

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

Hagerty, S. L., Hutchison, K. E., Lowry, C. A., & Bryan, A. D. (2020). An empirically derived method for measuring human gut microbiome alpha diversity: Demonstrated utility in predicting healthrelated outcomes among a human clinical sample. *PLoS ONE*, *15*(3). https://doi.org/10.1371/journal.pone.0229204

Hogan, I., Doherty, M., Fagan, J., Kennedy, E., Conneely, M., Brady, P., Ryan, C., & Lorenz, I. (2015). Comparison of rapid laboratory tests for failure of passive transfer in the bovine. *Irish*

Veterinary Journal, *68*(1), 1–10. https://doi.org/10.1186/s13620-015-0047-0

Holman, D. B., McAllister, T. A., Topp, E., Wright, A. D. G., & Alexander, T. W. (2015a). The nasopharyngeal microbiota of feedlot cattle that develop bovine respiratory disease. *Veterinary Microbiology*, *180*(1–2), 90–95. https://doi.org/10.1016/j.vetmic.2015.07.031

Holman, D. B., Timsit, E., & Alexander, T. W. (2015b). The nasopharyngeal microbiota of feedlot cattle. *Scientific Reports*, *5*. https://doi.org/10.1038/srep15557

Holman, D. B., Timsit, E., Amat, S., Abbott, D. W., Buret, A. G., & Alexander, T. W. (2017). The nasopharyngeal microbiota of beef cattle before and after transport to a feedlot. *BMC Microbiology*, *17*(1), 1–12. https://doi.org/10.1186/s12866-017-0978-6

Holman, D. B., Yang, W., & Alexander, T. W. (2019). Antibiotic treatment in feedlot cattle: A longitudinal study of the effect of oxytetracycline and tulathromycin on the fecal and nasopharyngeal microbiota. *Microbiome*, *7*(1), 86. https://doi.org/10.1186/s40168-019-0696-4

Hylander, B. L., & Repasky, E. A. (2019). Temperature as a modulator of the gut microbiome: what are the implications and opportunities for thermal medicine? *International Journal of Hyperthermia*, *36*(sup1), 83–89. https://doi.org/10.1080/02656736.2019.1647356

Idoate, I., Vander Ley, B., Schultz, L., & Heller, M. (2015). Acute phase proteins in naturally occurring respiratory disease of feedlot cattle. *Veterinary Immunology and Immunopathology*, *163*(3–4), 221–226. https://doi.org/10.1016/j.vetimm.2014.12.006

Ives, S. E., & Richeson, J. T. (2015). Use of Antimicrobial Metaphylaxis for the Control of Bovine Respiratory Disease in High-Risk Cattle. In *Veterinary Clinics of North America - Food Animal Practice*, *31*(3),341–350. W.B. Saunders. https://doi.org/10.1016/j.cvfa.2015.05.008

Janda, J. M., & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. In *Journal of Clinical Microbiology*, *45*(9), 2761–2764. American Society for Microbiology Journals. https://doi.org/10.1128/JCM.01228-07

Jerome, K. R. (2010). *Lennette's Laboratory Diagnosis of Viral Infections Fourth Edition InfectIous DIsease anD therapy serIes*.

Kimura, K., Reinhardt, T. A., & Goff, J. P. (2006). Parturition and hypocalcemia blunts calcium signals in immune cells of dairy cattle. *Journal of Dairy Science*, *89*(7), 2588–2595. https://doi.org/10.3168/jds.S0022-0302(06)72335-9

Kishimoto, M., Tsuchiaka, S., Rahpaya, S. S., Hasebe, A., Otsu, K., Sugimura, S., Kobayashi, S., Komatsu, N., Nagai, M., Omatsu, T., Naoi, Y., Sano, K., Okazaki-Terashima, S., Oba, M., Katayama, Y., Sato, R., Asai, T., & Mizutani, T. (2017). Development of a one-run real-time PCR detection system for pathogens associated with bovine respiratory disease complex. *Journal of Veterinary Medical Science*, *79*(3), 517–523. https://doi.org/10.1292/jvms.16-0489

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. *Journal of Clinical Microbiology*, *52*(2), 438–448. https://doi.org/10.1128/JCM.02485-13

Knapp, S., Winter, M., & Klotz, S. (2017). Increasing species richness but decreasing phylogenetic richness and divergence over a 320-year period of urbanization. *Journal of Applied Ecology*, *54*(4), 1152–1160. https://doi.org/10.1111/1365-2664.12826

Konopka, A., Lindemann, S., & Fredrickson, J. (2015). Dynamics in microbial communities: unraveling mechanisms to identify principles. *The ISME Journal*, *9*, 1488–1495. https://doi.org/10.1038/ismej.2014.251

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Larson, R. L., & Step, D. L. (2012). Evidence-Based Effectiveness of Vaccination Against Mannheimia haemolytica, Pasteurella multocida, and Histophilus somni in Feedlot Cattle for Mitigating the Incidence and Effect of Bovine Respiratory Disease Complex. *VFP*, *28*, 97-106.e7. https://doi.org/10.1016/j.cvfa.2011.12.005

Leigh, M. B., Pellizari, V. H., Uhlík, O., Sutka, R., Rodrigues, J., Ostrom, N. E., Zhou, J., & Tiedje, J. M. (2007). Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). *The ISME Journal*, *1*, 134–148. https://doi.org/10.1038/ismej.2007.26

Lima, S. F., Teixeira, A. G. V., Higgins, C. H., Lima, F. S., & Bicalho, R. C. (2016). The upper respiratory tract microbiome and its potential role in bovine respiratory disease and otitis media. *Scientific Reports*, *6*(1), 1–12. https://doi.org/10.1038/srep29050

Loong, T. W. (2003). Understanding sensitivity and specificity with the right side of the brain. In *British Medical Journal*, *327*(7417), 716–719. BMJ Publishing Group. https://doi.org/10.1136/bmj.327.7417.716

Lozupone, C., & Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Mahony, T. J., Paul, D., & Horwood, F. (2007). *Rapid detection of bovine respiratory disease pathogens FEEDLOTS*.

McDaneld, T. G., Kuehn, L. A., & Keele, J. W. (2018). Evaluating the microbiome of two sampling locations in the nasal cavity of cattle with bovine respiratory disease complex (BRDC). *Journal of Animal Science*, *96*(4), 1281–1287. https://doi.org/10.1093/jas/sky032

McMullen, C., Alexander, T. W., Léguillette, R., Workentine, M., & Timsit, E. (2020). Topography of the respiratory tract bacterial microbiota in cattle. *Microbiome*, *8*(1). https://doi.org/10.1186/s40168-020-00869-y

McMullen, C., Orsel, K., Alexander, T. W., van der Meer, F., Plastow, G., & Timsit, E. (2019). Comparison of the nasopharyngeal bacterial microbiota of beef calves raised without the use of antimicrobials between healthy calves and those diagnosed with bovine respiratory disease. *Veterinary Microbiology*, *231*, 56–62. https://doi.org/10.1016/j.vetmic.2019.02.030

Mizrahi-Man, O., Davenport, E. R., & Gilad, Y. (2013). Taxonomic Classification of Bacterial 16S rRNA Genes Using Short Sequencing Reads: Evaluation of Effective Study Designs. *PLoS ONE*, *8*(1), e53608. https://doi.org/10.1371/journal.pone.0053608

Moore, R. E., Kirwan, J., Doherty, M. K., & Whitfield, P. D. (2007). Biomarker discovery in animal health and disease: the application of post-genomic technologies. *Biomarker Insights*, *2*, 185–196. https://pubmed.ncbi.nlm.nih.gov/19662203/

Mosier, D. (2014). Review of BRD pathogenesis: The old and the new. *Animal Health Research Reviews*, *24*(6), 166–169. https://doi.org/10.1017/S1466252314000176

Muggli-Cockett, N. E., Cundiff, L. V., & Gregory, K. E. (1992). Genetic analysis of bovine respiratory disease in beef calves during the first year of life1. *Journal of Animal Science*, *70*(7), 2013–2019. https://doi.org/10.2527/1992.7072013x

NAHMS. (2013). National Animal Health Monitoring System. Feedlot 2011. Part IV: Health and health management on U.S feedlots with a capacity of 1000 or more head. In *USDA-APHIS-*

VS-CEAH-NAHMS (Issue March).

NRC. (2001). Dry matter intake. In: Nutrient Requirements of Dairy Cattle. In *Nutrient Requirements of Dairy Cattle* (7th ed.). National Academies Press. https://doi.org/10.17226/9825

Nguyen, N. H., Smith, D., Peay, K., & Kennedy, P. (2015). Parsing ecological signal from noise in next generation amplicon sequencing. *New Phytologist*, *205*(4), 1389–1393. https://doi.org/10.1111/nph.12923

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P. R., O'hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Szoecs, E., & Maintainer, H. W. (2019). *Package "vegan" Title Community Ecology Package Version 2.5-6*.

Ollivett, T. L., Caswell, J. L., Nydam, D. V., Duffield, T., Leslie, K. E., Hewson, J., & Kelton, D. (2015). Thoracic Ultrasonography and Bronchoalveolar Lavage Fluid Analysis in Holstein Calves with Subclinical Lung Lesions. *Journal of Veterinary Internal Medicine*, *29*(6), 1728–1734. https://doi.org/10.1111/jvim.13605

Pansri, P., Katholm, J., Krogh, K. M., Aagaard, A. K., Schmidt, L. M. B., Kudirkiene, E., Larsen, L. E., & Olsen, J. E. (2020). Evaluation of novel multiplex qPCR assays for diagnosis of pathogens associated with the bovine respiratory disease complex. *Veterinary Journal*, *25*(6), 105425. https://doi.org/10.1016/j.tvjl.2020.105425

Pardon, B., & Buczinski, S. (2020). Bovine Respiratory Disease Diagnosis: What Progress Has Been Made in Infectious Diagnosis? *In Veterinary Clinics of North America - Food Animal Practice, (36)*, 2, pp. 425–444. W.B. Saunders. https://doi.org/10.1016/j.cvfa.2020.03.005

Parker, A. M., Sheehy, P. A., Hazelton, M. S., Bosward, K. L., & House, J. K. (2018). A review of mycoplasma diagnostics in cattle. *Journal of Veterinary Internal Medicine*, *32*(3), 1241–1252. https://doi.org/10.1111/jvim.15135

Pielou, E. C. (1966). The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology*, *13*(C), 131–144. https://doi.org/10.1016/0022-5193(66)90013-0

Pimm, S. L. (1984). The complexity and stability of ecosystems. *Nature*, *307*(5949), 321–326. https://doi.org/10.1038/307321a0

Rappé, M. S., & Giovannoni, S. J. (2003). The Uncultured Microbial Majority. *Annual Review of Microbiology*, *57*(1), 369–394. https://doi.org/10.1146/annurev.micro.57.030502.090759

Rice, J. A., Carrasco-Medina, L., Hodgins, D. C., & Shewen, P. E. (2008). Mannheimia haemolytica and bovine respiratory disease. In *Animal Health Research Reviews*, *8*(2), 117–128. Anim Health Res Rev. https://doi.org/10.1017/S1466252307001375

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., & Firth, D. (2021). *Package "MASS."* http://www.stats.ox.ac.uk/pub/MASS4/

Rosengarten, R., Citti, C., Glew, M., Lischewski, A., Droeße, M., Much, P., Winner, F., Brank, M., & Spergser, J. (2000). Host-pathogen interactions in mycoplasma pathogenesis: Virulence and survival strategies of minimalist prokaryotes. *International Journal of Medical Microbiology*, *290*(1), 15–25. https://doi.org/10.1016/S1438-4221(00)80099-5

Roy, J. H. B. (1980). SYMPOSIUM: DISEASE PREVENTION IN CALVES Factors Affecting Susceptibility of Calves to Disease. *Journal of Dairy Science*, *63*(4), 650–664. https://doi.org/10.3168/jds.s0022-0302(80)82987-0

Sachse, K., Salam, H. S. H., Diller, R., Schubert, E., Hoffmann, B., & Hotzel, H. (2010). Use of a novel real-time PCR technique to monitor and quantitate Mycoplasma bovis infection in cattle herds with mastitis and respiratory disease. *Veterinary Journal*, *186*(3), 299–303. https://doi.org/10.1016/j.tvjl.2009.10.008

Schaefer, A. L., Cook, N. J., Bench, C., Chabot, J. B., Colyn, J., Liu, T., Okine, E. K., Stewart, M., & Webster, J. R. (2012). The non-invasive and automated detection of bovine respiratory disease onset in receiver calves using infrared thermography. *Research in Veterinary Science*, *93*(2), 928–935. https://doi.org/10.1016/j.rvsc.2011.09.021

Schaefer, A. L., Cook, N. J., Church, J. S., Basarab, J., Perry, B., Miller, C., & Tong, A. K. W. (2007). The use of infrared thermography as an early indicator of bovine respiratory disease complex in calves. *Research in Veterinary Science*, *83*(3), 376–384. https://doi.org/10.1016/j.rvsc.2007.01.008

Seker, E., Kuyucuoglu, Y., & Konak, S. (2009). Bacterial examinations in the nasal cavity of apparently healthy and unhealthy holstein cattle. *Journal of Animal and Veterinary Advances*, *8*(11), 2355–2359.

Sepulveda, J., & Moeller, A. H. (2020). The Effects of Temperature on Animal Gut Microbiomes. In *Frontiers in Microbiology*, *11*, 384. Frontiers Media S.A. https://doi.org/10.3389/fmicb.2020.00384

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., Højsgaardm, S., Fox, J.,

Lawrence, M. A., Mertens, U., Love, J., Lenth, R., & Bojesen Christensen, R. H. (2021). *Package "afex" Title Analysis of Factorial Experiments*.

Snowder, G. D., Van Vleck, L. D., Cundiff, L. V., & Bennett, G. L. (2006). Bovine respiratory disease in feedlot cattle: Environmental, genetic, and economic factors. *Journal of Animal Science*, *84*(8), 1999–2008. https://doi.org/10.2527/jas.2006-046

Sontakke, S., Cadenas, M. B., Maggi, R. G., Diniz, P. P. V. P., & Breitschwerdt, E. B. (2009). Use of broad range16S rDNA PCR in clinical microbiology. In *Journal of Microbiological Methods*, *76*(3),217–225. J Microbiol Methods. https://doi.org/10.1016/j.mimet.2008.11.002

Taylor, J. D., Fulton, R. W., Lehenbauer, T. W., Step, D. L., & Confer, A. W. (2010). The epidemiology of bovine respiratory disease: What is the evidence for predisposing factors? *Canadian Veterinary Journal*, *51*(10), 1095–1102.

Thanthrige-Don, N., Lung, O., Furukawa-Stoffer, T., Buchanan, C., Joseph, T., Godson, D. L., Gilleard, J., Alexander, T., & Ambagala, A. (2018). A novel multiplex PCR-electronic microarray assay for rapid and simultaneous detection of bovine respiratory and enteric pathogens. *Journal of Virological Methods*, *261*, 51–62. https://doi.org/10.1016/j.jviromet.2018.08.010

Thomas, A. C., Bailey, M., Lee, M. R. F., Mead, A., Morales-Aza, B., Reynolds, R., Vipond, B., Finn, A., & Eisler, M. C. (2019). Insights into Pasteurellaceae carriage dynamics in the nasal passages of healthy beef calves. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-48007-5

Tilman, D., Reich, P. B., Knops, J., Wedin, D., Mielke, T., & Lehman, C. (2001). Diversity and productivity in a long-term grassland experiment. *Science*, *294*(5543), 843–845. https://doi.org/10.1126/science.1060391

Timsit, E., Bareille, N., Seegers, H., Lehebel, A., & Assié, S. (2011). Visually undetected fever episodes in newly received beef bulls at a fattening operation: Occurrence, duration, and impact on performance. *Journal of Animal Science*, *89*(12), 4272–4280. https://doi.org/10.2527/jas.2011-3892

Timsit, Edouard, Hallewell, J., Booker, C., Tison, N., Amat, S., & Alexander, T. W. (2017). Prevalence and antimicrobial susceptibility of Mannheimia haemolytica, Pasteurella multocida, and Histophilus somni isolated from the lower respiratory tract of healthy feedlot cattle and those diagnosed with bovine respiratory disease. *Veterinary Microbiology*, *208*,

118–125. https://doi.org/10.1016/j.vetmic.2017.07.013

Timsit, Edouard, Holman, D. B., Hallewell, J., & Alexander, T. W. (2016). The nasopharyngeal microbiota in feedlot cattle and its role in respiratory health. *Animal Frontiers*, *6*(2), 44–50. https://doi.org/10.2527/af.2016-0022

Timsit, Edouard, Mcmullen, C., Amat, S., & Alexander, T. W. (2020). Respiratory Bacterial Microbiota in Cattle From Development to Modulation to Enhance Respiratory Health. *Veterinary Clinics of NA: Food Animal Practice*, *36*, 297–320. https://doi.org/10.1016/j.cvfa.2020.03.001

Timsit, Edouard, Workentine, M., van der Meer, F., & Alexander, T. (2018). Distinct bacterial metacommunities inhabit the upper and lower respiratory tracts of healthy feedlot cattle and those diagnosed with bronchopneumonia. *Veterinary Microbiology*, *221*, 105–113. https://doi.org/10.1016/j.vetmic.2018.06.007

Tyler, J. W., Hancock, D. D., Thorne, J. G., Gay, C. C., & Gay, J. M. (1999). Partitioning the Mortality Risk Associated with Inadequate Passive Transfer of Colostral Immunoglobulins in Dairy Calves. In *J Vet Intern Med* (Vol. 13).

Uehara, A., Fujimoto, Y., Fukase, K., & Takada, H. (2007). Various human epithelial cells express functional Toll-like receptors, NOD1 and NOD2 to produce anti-microbial peptides, but not proinflammatory cytokines. *Molecular Immunology*, *44*(12), 3100–3111. https://doi.org/10.1016/j.molimm.2007.02.007

USDA. (2014). Dairy 2014 Health and Management Practices on U.S. Dairy Operations. *Dairy 2014 Health and Management Practices on U.S. Dairy Operations. 2014*.

Vogel, G. J., Bokenkroger, C. D., Rutten-Ramos, S. C., & Bargen, J. L. (2015). A retrospective evaluation of animal mortality in US feedlots: rate , timing , and cause of death. *Bovine Practitioner*, *49*(2), 113–123. https://doi.org/10.21423/bovine-vol49no2p113-123

Waters, C. M., & Bassler, B. L. (2005). Quorum sensing: Cell-to-cell communication in bacteria. In *Annual Review of Cell and Developmental Biology*, *21*, 319–346. Annu Rev Cell Dev Biol. https://doi.org/10.1146/annurev.cellbio.21.012704.131001

Watts, J. L., & Sweeney, M. T. (2010). Antimicrobial Resistance in Bovine Respiratory Disease Pathogens: Measures, Trends, and Impact on Efficacy. *Veterinary Clinics of North America - Food Animal Practice*, *26*(1), 79–88. https://doi.org/10.1016/j.cvfa.2009.10.009

White, B. J., & Renter, D. G. (2009). Bayesian estimation of the performance of using clinical

observations and harvest lung lesions for diagnosing bovine respiratory disease in post-weaned beef calves. In *Journal of Veterinary Diagnostic Investigation*, *21*(4), 446–453. https://doi.org/10.1177/104063870902100405

Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, *30*(3), 279–338. https://doi.org/10.2307/1943563

Wildman, B. K., Perrett, T., Abutarbush, S. M., Timothy Guichon, P., Pittman, T. J., Booker, C. W., Schunicht, O. C., Kent Fenton, R., & Kee Jim, G. (2008). A comparison of 2 vaccination programs in feedlot calves at ultra-high risk of developing undifferentiated fever/bovine respiratory disease. In *CVJ* (Vol. 49).

Wilson, B. K., Richards, C. J., Step, D. L., & Krehbiel, C. R. (2017). Beef species symposium: Best management practices for newly weaned calves for improved health and well-being. *Journal of Animal Science*, *95*(5), 2170–2182. https://doi.org/10.2527/jas.2016.1006

Wolfger, B., Timsit, E., White, B. J., & Orsel, K. (2015). A Systematic Review of Bovine Respiratory Disease Diagnosis Focused on Diagnostic Confirmation, Early Detection, and Prediction of Unfavorable Outcomes in Feedlot Cattle. *Veterinary Clinics of North America - Food Animal Practice*, *31*(3), 351–365. https://doi.org/10.1016/j.cvfa.2015.05.005

Zeineldin, M., Lowe, J., & Aldridge, B. (2020a). Effects of Tilmicosin Treatment on the Nasopharyngeal Microbiota of Feedlot Cattle With Respiratory Disease During the First Week of Clinical Recovery. *Frontiers in Veterinary Science*, *7*. https://doi.org/10.3389/fvets.2020.00115

Zeineldin, M. M., Lowe, J. F., Grimmer, E. D., De Godoy, M. R. C., Ghanem, M. M., Abd El-Raof, Y. M., & Aldridge, B. M. (2017). Relationship between nasopharyngeal and bronchoalveolar microbial communities in clinically healthy feedlot cattle. *BMC Microbiology*, *17*(1). https://doi.org/10.1186/s12866-017-1042-2

Zeineldin, M., A Elolimy, A., & Barakat, R. (2020b). Meta-analysis of bovine respiratory microbiota: link between respiratory microbiota and bovine respiratory health. FEMS Microbiology Ecology, 96(8), 127. https://doi.org/10.1093/femsec/fiaa127

Zeineldin, M., Lowe, J., & Aldridge, B. (2019). Contribution of the Mucosal Microbiota to Bovine Respiratory Health. In *Trends in Microbiology*, *27*(9), 753–770. Elsevier Ltd. https://doi.org/10.1016/j.tim.2019.04.005

Zhang, S., Chen, D. C., & Chen, L. M. (2019). Facing a new challenge: The adverse effects of

antibiotics on gut microbiota and host immunity. In *Chinese Medical Journal*, *132*(10), 1135–1138. Lippincott Williams and Wilkins. https://doi.org/10.1097/CM9.0000000000000245