APPLIED BACTERIAL ECOLOGY IN LIVESTOCK SYSTEMS

by

Carmen L. Wickware

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Animal Sciences

West Lafayette, Indiana

December 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Timothy A. Johnson, Chair

Department of Animal Sciences

Dr. Cindy Nakatsu

Department of Agronomy

Dr. Darryl Ragland

Department of Veterinary Clinical Sciences

Dr. Jacquelyn Boerman

Department of Animal Sciences

Dr. Paul Ebner

Department of Animal Sciences

Approved by:

Dr. Kolapo Ajuwon

To my family, who have supported me in every way. I love you all dearly.

ACKNOWLEDGEMENTS

To my advisor, Dr. Tim Johnson. I am so glad I decided to talk to you after the conference. I have gained so much knowledge and have been given so many different opportunities since being a member of your lab. Thank you for trusting me to be your first graduate student. Thank you for allowing me to stretch my skills of mentoring with so many students and researchers. Just, thank you.

To my committee, Drs. Cindy Nakatsu, Daryll Ragland, Jackie Boerman, and Paul Ebner. You have all challenged me to think about my research outside of the individual settings. I appreciate you all for your time and expertise.

To Ulyssa, Samira, Torrie, Emily, Vanessa who have taught me to be unapologetically myself. I am glad to know you and love you all very much.

To Laura, Cat, and Giselle who have continued to be my friend after all these years. Your support means the world to me.

To Mgavi, whose constant mentoring and kind words have set me up for success in ways that are beyond measure.

To the Drs. Wickware who have inspired me from the beginning. I have joined the ranks and hope I can do the name justice.

To my parents, without whom I would literally not be here today. I have grown so much, but I will always be your child.

And finally, to Chris. You are my forever Pard. Travelling through Byrgenwerth just wouldn't be the same without you and Ein by my side.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Microbiome studies are varied and involve the examination of microorganisms at different levels: individual cells to determine individual functions, populations of specific microorganisms to determine interactions between organisms, and/or communities of microorganisms for a broader investigation of interactions between organism and environment. These studies are typically done within the context of a particular niche or environment. There are two parts to this dissertation, separated by the types of research involved. First, the analysis of bacterial communities using 16S rRNA sequencing and analysis. In this first part the bacterial communities of the reproductive tract of bulls and the gastrointestinal tract of weanling pigs were studied. The reproductive organs of the male, domestic species had not been studied from an ecological perspective prior to the study. As such, the research was mainly focused on characterizing the bacterial communities found within the prepuce of bulls that were considered to be healthy, or that the breeding soundness exam was satisfactory and the bulls had no clinical disease in the urogenital tract. Through this study two distinct types of bacterial communities were found based on the diversity of the observed taxa; the groups were split into a low diversity group identified by the presence of *Bradyrhizobium* and a high diversity group distinguished by the abundance of mucosal-associated bacteria found in oral, respiratory, and vaginal communities of cattle. Second, the effects of supplementary, soluble fiber on the intestinal bacterial communities of piglets pre- and/or post-weaning were studied. The rationale behind this study was to determine if pre-weaning fiber could alter the microbiome prior to weaning and the change of diet from liquid to solid. Pre-weaning, supplementary, soluble fiber was found to increase short-chain fatty acid concentrations and bacterial taxa potentially involved in their production. Additionally, bacterial taxa implicated in an increased inflammatory response were reduced in groups fed supplementary fiber. Taken together, the two bacterial community studies highlight the gaps in knowledge for reproductive communities in male animals as well as the potential for reducing weaning stress in pigs. Part two of this dissertation focuses on whole genome sequence analysis as a way to study bacterial populations associated with bovine respiratory disease (BRD), a common and potentially fatal disease in cattle. Identification of BRD has low accuracy and the presence of antibiotic resistant bacteria increases the chance of treatment failure. Using machine learning, the prediction of antibiotic resistance in bacterial isolates from animals with BRD was performed to find potential sequences for use in future molecular assays.

While using known resistance genes was helpful for some antibiotics, several of the antibiotics used in treating BRD were better predicted using the machine learning models. Model output sequences should be further tested using molecular methods to determine function and importance before using as an assay target. Put together, the contents of this dissertation should serve as an introduction to bacterial ecology as well as how the concepts can be applied to food animal production systems.

CHAPTER 1. INTRODUCTION TO BACTERIAL ECOLOGY AND ITS RELEVANCE TO LIVESTOCK PRODUCTION

Abstract

Animal agriculture is one of the largest economic sectors globally. With the growing, global human population, animal producers have been challenged with improving the efficiency of animal growth while maintaining animal health and welfare. Research into nutrition and disease mitigation has increased as scientists try to find ways of increasing final product quantity and quality through optimizing animal feed efficiency (i.e., decreasing the amount of feed needed to produce an equal or greater amount of meat, milk, eggs) and increasing animal resilience to disease. Microbial communities have become one of the main targets for animal research as studies have shown the importance of studying the various organs and systems of animals (e.g., lungs, reproductive tract, gastrointestinal tract, etc.). Bacteria aide in the digestion of feed into smaller components that are absorbed and used by the body. Additionally, commensal bacteria can help provide a preventative barrier against pathogenic microorganisms. Pathogenic bacteria can invade epithelial cells with some species able to avoid the host animal immune system through proteins expressed on their surface or secreted extracellularly. Through the study of bacterial communities, populations, and cells we can improve our understanding of the ways bacteria interact with other microorganisms, host environments, and feed as a component of productivity animal health.

1.1 Background

Animals raised for the production of meat, eggs, or milk for consumption have been a widely studied system as the global population continues to rise. Meat, dairy, and eggs have traditionally been the most consumed sources of protein, globally (FAO, 2018). While the role microbial communities play in human health have been studied extensively, there is increasing motivation to do the same in livestock. Livestock producers and scientists have started relying on the contribution of the animal's microbial communities as a way to increase yield of animal products while trying to maintain high health status of the animals (Donham, 2000; Sundrum, 2015; Temple & Manteca, 2020).

Recent advances in microbial analysis (i.e., community analysis, whole genome sequence analysis, etc.) have led to increased attention into the relationships between microorganisms, nutrition, and disease resistance in livestock production (Dewi & Kollanoor Johny, 2022; Kogut & Arsenault, 2016; Y. Li et al., 2020; Monteiro et al., 2022; O'Hara et al., 2020). Bacterial communities can be studied to find sources of beneficial bacteria, and examining the various aspects of pathogenic bacteria, such as the most common pathogens for a disease or environmental factors increasing risk of bacterial disease, to help improve animal health. While the microbial communities and populations in the gastrointestinal tract (GIT) of animals have significant examination, there is increased awareness of the potential impacts microorganisms can have on other organs and systems within livestock.

Naturally, as more is learned about the complex interactions between microorganisms and animals, the ways in which we approach microbial ecology research shift. In fact, some scientists think of the gut microbiome as a separate organ due to the ability of microorganisms to manipulate metabolic and immune processes in the host animal (Anwar et al., 2020). The microbiome can be heritable based on parental genetics, and changes in the humoral immunity bring about small, "microevolutionary" shifts in gut microbial community members (L. Yang et al., 2017). The intricate connections between microorganisms and host create ecological niches and require individual examination to determine the differences within each community.

One of the most consistent findings in animal bacterial community studies is the age-related shift in intestinal bacterial communities. Attempts to invoke early maturation of the communities in animals have been seen across species as a way to reduce risk of disease and increase animal performance (Awad et al., 2016; Choudhury et al., 2021; Fu et al., 2022). As animals mature, the alpha diversity of bacterial communities increases in richness (number of observable or predicted bacterial taxa) and evenness (the distribution of abundance of each taxon) when compared to younger animals (Guo et al., 2020). Additionally, a more stable diet selects specific taxa of bacteria, thus the phylogenetic diversity becomes more consistent through the presence of certain bacterial taxa as mentioned above. Aruwa et al. (2021) highlighted the ways in which the poultry GIT is impacted during different stages of life. First, initial colonization of bacteria affects disease resistance. Colonization by pathogenic bacteria leads to poor health early, whereas symbiotic or commensal bacteria colonization increases disease resistance by actively defending against

pathogen colonization (Aruwa et al., 2021). Second, successional changes in the poultry gut microbial community eventually stabilize in the mature animal.

There are many types of studies in microbial ecology, but the focus of this dissertation will be on bacterial community and population level analysis, referencing larger microbial ecology topics where needed. The remainder of this chapter serves as an introduction into some basic aspects of bacterial ecology, how the concepts can be applied to livestock production systems, and the ways in which we use bioinformatic analysis on bacterial communities and populations from livestock to help improve animal health and performance.

1.2 Bacterial ecology

Microbial communities are defined as a group of microorganisms living and interacting with each other in a specific niche under similar environmental conditions. These communities can be a mixture of all types of microorganisms – viruses, fungi, bacteria, and protozoa. The rationale behind studying microbial communities is to determine their interactions with biotic and abiotic components, i.e., host, environment, and other microorganisms. These interactions make up the basis of biological activity as microorganisms. There are an estimated 39 trillion microbial cells in one human body and bacterial cells alone make up 30% of those cells, contributing to about 15% of the total cells within the human body (Gilbert et al., 2018). Bacteria commonly provide beneficial functions such as metabolizing complex carbohydrates and fats into smaller components, more readily available for the body to absorb as well as producing metabolites such as essential amino acids and vitamins (Albenberg & Wu, 2014; Dieterich et al., 2018).

In addition to providing nutritional benefits, some bacteria function as a physical and chemical barrier on mucosal surfaces to prevent pathogenic microorganisms from attaching and invading (Paone & Cani, 2020). Through competitive exclusion, commensal bacteria can physically prevent pathogens from attaching to surfaces or utilize nutrients necessary for the growth of pathogens(Hibbing et al., 2010). The mucosal lining of many organs with outward-facing cells (lungs, intestines, etc.) helps to provide a niche wherein established microbiota can provide additional defenses. Production of antimicrobial compounds, by both the host and mucosal-associated microorganisms, grant mucosal-associated bacteria a protective gradient through which pathogenic microorganisms must travel to invade host cells (Paone & Cani, 2020; Perez-Lopez et al., 2016).

Studying bacterial communities allows a broad picture of the potential functions in the animal environment, such as the bacterial community of the intestinal mucosa described previously. However, examining single populations of specific bacteria within a community allows the determination of individual benefits and detriments of each member of the community. Both types of research are important and expand the knowledge of how bacteria interact with the host and the environment. Their individual importance is discussed in the next several paragraphs.

Increasing numbers of biological functions have been linked to the activity of microorganisms. Studies by Estrela et al. (2021) and Pacheco et al. (2021) showed the link between nutrient source or environment and microbial community members, with family-level changes being fairly predictable with different nutrients, but individual genera of the family changing. Though originally for human research, examination of the interconnection between the gut and brain in animals has displayed the ways in which microbial interactions, including commensal, symbiotic, or pathogenic microorganisms, affect behavior and appetite (Diaz Heijtz et al., 2011; Kraimi et al., 2019; Suchodolski, 2018).

On the other hand, the analysis of bacterial isolates allows the connection of phenotypic and genotypic functions. As described above, potential functions can be determined by looking through the genome to identify coding sequences linked to a function. However, coding sequences do not always get translated, due to mutations in the coding sequence (i.e., deletion, insertion, frameshift), weak or absent promoter sequences, or the requirement of specific environmental conditions (i.e., heat, pH, and oxygen stress). By analyzing a specific phenotype in conjunction with the functional information gathered from the genome, the ability of a specific organism to perform a particular function can be elucidated.

To determine the contributions of individual bacteria to the community, it is important to study specific populations of bacteria (e.g., *Lactobacillus* spp., *Escherichia coli*, *Salmonella enterica* ssp.). Examining explicit bacterial functions of individual commensal and pathogenic populations can help define the limits of a bacterial population such as how specific species are selected within an environment and how a specific population becomes virulent (Sheppard et al., 2018; Tzortzis et al., 2005; VanInsberghe et al., 2020; Y. Wu et al., 2021).

In addition to determining differences between populations, there is increased benefit to studying the change in a population over time (Bossi et al., 2003; Kamath et al., 2016; Ward et al., 2014). Studying the longitudinal changes of a population grants the study of genetic changes and

mechanisms for those changes. For example, Rangel et al. (2005) studied E. coli O157:H7 outbreaks from 1982-2002 and determined that the outbreaks continued to be foodborne with increased occurrence from water and animal contact cases. In various studies on the genetic disparity of E. coli O157:H7 strains, it was discovered there are strains with varying resistance to processing parameters such as high pressure, acidity, and temperature (Kay et al., 2017; Lim et al., 2010; Malone et al., 2006). The longitudinal studies of bacterial populations can increase our knowledge of factors that lead to virulence, as discussed above with E. coli O157:H7 but also the value of beneficial bacteria, such as Lactobacillus spp., a genus that has been studied extensively in humans for effects on digestion and resistance to pathogenic microorganisms (Heeney et al., 2018; Walter, 2008) and has gained the attention of animal producers (Dewi & Kollanoor Johny, 2022; Hu et al., 2018; Mountzouris et al., 2007; Valeriano et al., 2017).

Phenotypic resistance to antimicrobials, and antibiotics in particular, adds to the list of functions that have growing attention from scientists. The diminishing number of effective antibiotics has put an emphasis on determining antibiotic resistant bacteria (ARB) and the mechanisms by which bacteria become resistant (Bengtsson-Palme, 2018; Dowling et al., 2017). Pairing genomic antibiotic resistance data with observable resistance phenotype is important for the reasons mentioned above; the existence of a particular gene does not always result in transcription and translation. Scanning for particular antibiotic resistance genes (ARG) within a community is even less helpful when bacteria are intrinsically resistant to a particular antibiotic and may not contain the resistance genes despite being phenotypically resistant (Cox & Wright, 2013). This connection between genotype and phenotype is not always straightforward and requires the isolation of bacteria to perform biochemical and molecular assays (Leclercq et al., 2013). Linking antibiotic resistance (AR) to particular bacterial populations within a community is an obstacle requiring novel methods to be developed to quickly detect AR populations and improve treatment of diseases.

1.3 Increasing animal health by studying bacterial communities and populations

In livestock, microbial studies have been focused on the gut microbiome, however there is evidence across all animals of the importance of microbial communities on the skin, in the lungs, and other organs. Animal production as a whole includes much more than nutrition and disease treatment, and recent findings have encouraged the scientific community to study the whole animal

instead of concentrating efforts only on the GIT. Adnane & Chapwanya (2022) reviewed the importance of the microbial diversity in the reproductive tract of cattle, discussing the initial establishment of the reproductive microbiome early in life and the connection of microbiome dysbiosis to infertility and uterine diseases. Similarly, Sanglard et al. (2020) studied the impact of porcine reproductive and respiratory syndrome (PRRS) vaccination on the vaginal microbiome of sows and found the microbiome was different between sows with low and high reproductive performance. In both the reproductive studies, connections between microbial community members and animal performance or disease (e.g., infertility and number of progeny) were found, suggesting additional investigation is needed determine the importance of the reproductive tract microbiome on breeding programs.

In studies on the respiratory microbiome of livestock, key bacterial and viral agents have been identified for most animal species. Research on the swine respiratory microbiome shows *Mycoplasma hyopneumoniae* to be one of the primary bacterial pathogens in cases of swine pneumonia (Siqueira et al., 2017). In fact, *Mycoplasma* spp. are implicated in severe cases of respiratory disease in mammalian and avian livestock species (Daee et al., 2020; Dudek et al., 2020; Gerchman et al., 2011; Maes et al., 2020). *Pasteurella multocida*, another respiratory bacterial pathogen found in both mammalian and avian livestock species (Hsuan et al., 2009; Rawat et al., 2019; J.-R. Wu et al., 2003), is observed as an opportunistic pathogen, requiring a lowering of immune defense from viruses or stress before becoming virulent.

Stresses, including but not limited to transportation, temperature change, dietary change (e.g., from milk to solid feed), co-mingling, and weaning impact animals at every stage of life. Each of these stresses has been implicated in poor performance in young animals, increased prevalence of disease, and a lower quality of finished product (e.g., milk yield and carcass weight). Researchers have approached stress in various ways to reduce the negative effects on animals. A review on stress in early weaned piglets by Campbell et al. (2013) highlights weaning as one of the most stressful times in a pig's life. The authors consider separation from the sow, handling, establishment of new social hierarchies, and change in physical environment as some of the points of stress during weaning that lead to low performance and decreased health status of pigs later in life (Campbell et al., 2013).

The stress of changes to an environment can include physical occurrences such as movement to a different facility, but also factors such as temperature and humidity. Increases in

temperature are linked to a decline in gut integrity of animals with heat stress leading to impaired gut barrier function, an increase in intestinal inflammation. Prolonged exposure to high temperatures can permanently damage the intestinal epithelia (Koch et al., 2019; Quinteiro-Filho et al., 2012; Summer et al., 2019). As gut barrier function is reduced, there is a higher risk of infection due to the ability of bacteria and other microorganisms to bypass gut defenses and be disseminated into other parts of the body (Spadoni et al., 2015; Wickramasuriya et al., 2022).

However, it is important to note that the gut barrier does not only refer to the intestinal epithelia and immune cells but includes the microbiota in the lumen and mucosa as well. A study by Xia et al. from 2022 connected a reduction in gut barrier function to dysbiosis of the microbiome within the gut. The mucosal-associated bacteria are exposed to more oxygen than their luminal counterparts due to the oxygen needed within the gut epithelial cells being diffused through the mucous barrier. However, Xia et al. (2022) discussed the impact small changes in oxygen levels can have on the mucosal bacteria, with increased reactive oxygen species (ROS) being created as oxygen is moved to the peripheral vascular system during heat stress. The ROS, created by both the epithelial cells as well as the mucosal bacteria, cause increased damage to the epithelia as well as the microbiota within the mucous and lumen (Xia et al., 2022).

While weaning stress greatly affects the function of the GIT as discussed above, many other essential systems are disturbed by maternal separation and weaning. Malmuthuge et al. (2021) studied the separation of calves from dams and subsequent transportation to determine the effects of these stressors on the immune system of the young animals. The authors found that while the microbiome of the calves' upper respiratory tract did not significantly change compared to suckling calves, the expression of neuroimmune receptor genes increased, warranting further investigation into the immune responses of weaned calves (Malmuthuge et al., 2021). Studies in early-weaned piglets show the animals often benefit from supplementation of amino acids (J. Wang et al., 2021), probiotics (Xu et al., 2018), or prebiotics (Hu et al., 2018) as ways to ameliorate the effects of weaning stress. In a study by Choudhury et al. (2021), an attempt was made to mature the gut microbial community of piglets to reduce weaning stress and prevent post-weaning diarrhea. Using a diet high in fiber during farrowing, the authors observed a change in the gut microbiota, increased weight gain, and decreased diarrhea scores post-weaning compared to piglets without access to the solid feed (Choudhury et al., 2021).

Stress to livestock can also be the result of poor air and water quality. Studies have examined the microbial communities of the environments within and surrounding production facilities to determine the impact of air and water quality on animal health from a microbial perspective. In a study on the air quality within calf barns during winter, Lago et al. (2006) cultured airborne bacteria to determine which factors can result in lower bacterial counts, as poor ventilation can increase the risk of respiratory disease. The authors found factors such as pen bedding type, temperature, ventilation rate, and openings between pens contributed to the number of bacteria within individual pens as well as in the shared space (alley) in the barns (Lago et al., 2006). Ventilation type also contributes to differences in respiratory disease as ventilation helps control the temperature and humidity within animal housing as well as the removal of ammonia and carbon dioxide. In a 2022 study by Shi et al., the authors observed tunnel ventilation was the best option for cattle during winter as it takes advantage of higher temperature within the soil compared to open air.

1.4 Nutrition and bacterial communities

The microbial environment of an animal's GIT plays a distinct role in health status and production capabilities. Without gut microbial community members able to degrade complex carbohydrates, proteins, and fats into smaller components, animals will not have the ability to incorporate the components into muscle, milk, or eggs (Hooper et al., 2002; Wessels, 2022). While the microbial community can affect how well an animal can absorb nutrients from their diet, the reverse relationship is also true – animal feed can impact the microbial community within the GIT, selecting for microorganisms that can efficiently utilize the dietary components. For this reason, the most suitable diet to balance nutrition for both animals and beneficial microorganisms while deterring the colonization and growth of pathogenic microorganisms is under current investigation.

Optimizing absorption of nutrients and reducing nutrient loss through waste are key components to increasing feed efficiency. Utilization of dietary components (e.g., fats, protein, carbohydrates, and vitamins) have been studied through full diet changes as well as supplementation as ways to increase feed efficiency and reduce unwanted excretion (de Vries et al., 2015; Heinritz et al., 2016). The gut bacterial community can improve the full utilization of nutrients by producing metabolites and enzymes to help process complex carbohydrates (H. Wang

et al., 2020) and fats (Kim et al., 2020; H. Yang et al., 2018) into useable components (e.g., secondary metabolites) for the host animal.

Studies on the swine GIT have emphasized the importance of fiber in the piglet diet and how incrementally introducing solid feed to the suckling pig (creep feed) can influence the development of the piglet intestine later in life. Chen et al. (2021) studied the differences between types of creep feed (powdered, soft pellet, or hard pellet) on performance and intestinal development. The authors found that soft pelleted creep feed improved growth performance and intestinal development in piglets measured 10 days post-weaning. However, not all studies have a positive result from using creep feed (Middelkoop et al., 2020), suggesting the implementation potential or generalizable benefit is unknown or narrow.

The use of creep feed highlights one of the key findings of studies in livestock; dietary fiber is linked to the maturation of the GIT as well as the microbial community. The type of fiber used is dependent on species of animal, but is considered a crucial dietary component in monogastric (Jha et al., 2019) and ruminant (Mertens, 1997) animals for production of short chain fatty acids (SCFA). In both monogastric and ruminant animals, the production of SCFA is performed by microorganisms; in monogastric animals this happens primarily in the cecum and/or colon while in ruminant animals, production of SCFA occurs primarily in the rumen. Fiber is supplemented in the monogastric diet to select SCFA producing bacteria that are considered beneficial (Jha et al., 2019). The SCFA are utilized for many purposes including as energy for colonic epithelial cells, creating a better barrier and quicker responses of immune cells. Fiber is used in a similar way in ruminant animals, where selection of microbial communities with low numbers of unique but specialized microorganisms effective in degrading plant biomass have been implicated in dairy cattle with high feed efficiency (Shabat et al., 2016).

Secondary metabolites produced by bacteria can be used by the animal, like the SCFA mentioned above, but also may provide an environment that is selective for other microorganisms. In ruminants, lactic acid producing bacteria and exogenous lactic acid found in silages create an environment selective for lactic acid utilizing microorganisms, striking a balance between the fermentation of plant biomass, metabolism of lactic acid by the animal for energy, and utilization of excess lactic acid to prevent acidosis (Mackenzie, 1967; Mills et al., 2014). Production of urease by ureolytic bacteria in the animal GIT is important in the utilization of nitrogen and the prevention of excess urea in both ruminants and monogastric animals (Patra & Aschenbach, 2018). Finding a

stable microbial community that optimizes the utilization of nutrients and removal of waste between animal and microorganism is crucial to animal health and efficiency.

1.5 Animal disease treatment and AMR in bacterial communities

Thus far, we have discussed the ways in which microorganisms can benefit animals through microbial communities that enhance feed efficiency and improving animal health by balancing nutrient utilization and loss. However, pathogenic microorganisms are always part of the equation, and while a stable microbial community can help reduce the chances of infection, animals that become sick from a microorganism often demand the use of antimicrobials (Johnston, 1998). An antimicrobial is any compound that inhibits the growth of, or actively kills, microorganisms including antibiotics, disinfectants, and heavy metals.

Antimicrobials play a large role in the health and performance of animals (Broom, 2017; Butaye et al., 2003; Dibner & Richards, 2005).While producers vaccinate animals against common diseases caused by microorganisms, antimicrobials are also used for prophylactic treatment of diseases known to be a problem, e.g., diarrhea in grow-finish operations in (USDA - APHIS, 2002). In prophylactic treatment of disease, the dose of an antimicrobial given is typically subtherapeutic which is much lower than what would be used to treat a disease (Edrington et al., 2001; Van Cuong et al., 2021).

However, the use of antimicrobials at therapeutic and subtherapeutic levels can lead to an increase in antimicrobial resistance (AMR). Rosenfeld (2017) reviewed the effects of metal, antimicrobial, and other chemical contamination on the microbial communities within animals and concluded that gut dysbiosis is an acute result of exposure to anthropogenically contaminated water, but prolonged exposure can lead to systemic issues. Similar to water quality, diets are often supplemented with metals such as zinc and copper to increase feed intake, to provide essential micronutrients, or to provide prophylactic treatment of digestive dysbiosis (Carlson et al., 2007; Villagómez-Estrada et al., 2020). However, the antimicrobial properties of metal in water or feed can select for antimicrobial resistant bacteria within the animal gut (Baker-Austin et al., 2006; M. Zhang et al., 2017), leading to treatment complications.

Though commonly applied on farms for preventative and therapeutic practices, the use of antimicrobials is well-regulated. Historically, antimicrobials were also used as growth promoting compounds (AGP) in livestock (Broom, 2017; Butaye et al., 2003; Dibner & Richards, 2005).

Animals that receive growth promoters perform better than their un-treated counterparts, with increased disease tolerance and feed efficiency (Cromwell, 2002; Dibner & Richards, 2005; Shabat et al., 2016). However, recent regulation of antimicrobials has led to AGP being prohibited in several countries, including the United States (Department of Health and Human Resources, 2015). In the US, government agencies and organizations (e.g., Department of Agriculture (USDA), Food and Drug Administration (FDA), and the Animal Health Association (USAHA), as well as global regulatory agencies such as the World Health Organization (WHO) and the Food and Agriculture Organization (FAO), give regulations and guidelines, or encourage legislation that, in combination with veterinarians and producers, result in the best treatment for a sick animal (Food and Drug Administration, 2012; World Health Organization, 2019). Using the most judicious treatment is part of the global initiative in reducing the amount of antimicrobials used to increase the effectiveness of available antimicrobials (Forum on Microbial Threats et al., 2017).

The reduction in the use of commonly known antimicrobials has spurred research into antibiotic alternatives – feed additives, management practices, and biosecurity measures – that can help reduce the prevalence of disease and need for antimicrobials (Allen et al., 2013; Sharma et al., 2018; Suresh et al., 2018). While antimicrobials are useful in the treatment and prevention of disease, the search continues for alternative treatments for illnesses to reduce the need to use antimicrobials except in very particular cases of infection. The motivation behind finding alternatives to antimicrobials is the increased occurrence of antimicrobial resistant (AMR) microorganisms (Bonten et al., 2001; Forum on Microbial Threats et al., 2017; Sharma et al., 2018).

However, antimicrobial resistance in bacteria is not confined to pathogens. In fact, the interactions of commensal and pathogenic bacteria in disease progression has attracted the focus of scientists as it has been found that commensal bacterial populations are high in AMR (Card et al., 2017; Chuppava et al., 2019; Ramos et al., 2022; Tawfick et al., 2022). Populations of commensal *E. coli* in the GIT of humans and animals are reservoirs for resistance plasmids (Card et al., 2017; Chuppava et al., 2019; Tawfick et al., 2022) and AMR in commensal *Staphylococcus aureus* in wild animal populations have been connected to agricultural land usage (Ramos et al., 2022) suggesting the movement of AMR between environments.

Much like pathogens, commensal organisms are trying to survive (e.g., exposure to antimicrobials, environmental changes, nutrient scarcity) and can develop resistance genes to the

stressors. Commensal bacteria are commonly more abundant and persistent within an environment than pathogenic bacteria thus the probability of commensal organisms developing resistances is high (Lopetuso et al., 2013; McLaren & Callahan, 2020). Antimicrobial resistance in commensal bacteria can be a benefit to the host as they can serve as a barrier to pathogen colonization and invasion of cells (McLaren & Callahan, 2020). However, the transfer of antimicrobial resistance from commensal to the pathogenic bacteria poses a risk to animal and human health.

Resistance genes can move to pathogenic bacteria through horizontal gene transfer (HGT) which can occur through conjugation (direct exchange between bacteria), transduction (infection bacteriophage), or transformation. The mechanisms of HGT have been well studied and there are studies that have attempted to characterize the scenarios that make HGT more likely to occur, such as spatial ecology during exposure to low concentrations of antibiotics (Cairns et al., 2018) and the large diversity of microorganisms in the mammalian GIT (Shterzer & Mizrahi, 2015). The animal gut, including humans, has long been associated with high HGT of antibiotic resistance genes and thus has been the primary focus of animal researchers (Lima et al., 2020; Shterzer & Mizrahi, 2015). As HGT is thought of as the main contributor to pathogen-acquired resistance genes, it has gained attention from various disciplines (Cairns et al., 2018; Jutkina et al., 2016; Úbeda et al., 2005).

Overall, treating animal disease has the same end goal as with human clinical treatments – to reduce or eliminate the cause of the disease while quickly and safely returning the animal to a healthy state. There are additional intentions when treating a disease in agricultural animals that include lessening the impact of the treatment on surrounding environments (including soils, surface waters, etc.), animals (companion, production, or wild), and humans. However, the considerations in each of these cases involves making the best choice possible with the information at hand, making animal disease resilience and effective treatments research critical to reduce antibiotic resistance and to preserve the efficacy of antibiotics.

1.6 Determination of bacterial functions in communities and populations

Prior to the “golden age of sequencing” happening now, analysis of bacteria and bacterial disease was limited to what bacteria could be cultured and easily characterized by biochemical and/or molecular methods (Didelot et al., 2012; Schabereiter-Gurtner et al., 2001). Now, with the ability to go from isolation to fully sequenced bacterial genome in the matter of days, researchers

have put more emphasis on collecting vast libraries of isolates from which genomic information can be obtained. While the most genomic information is still obtained through culturing and sequencing of bacterial isolates, broad examination of bacterial communities within an environment can be performed with the advent of meta'omic (i.e., metagenomic, metatranscriptomic, etc.) sequencing and analysis (Segata et al., 2013).

The increasing interest in bacterial function and selection in livestock necessitates the analysis of large amounts of sequence data. Bioinformatics, a field growing in popularity among scientists of many disciplines, permits multifactor analysis of bacterial sequence data. Studying bacterial communities gives us insight into potential reservoirs of beneficial and detrimental processes for animals. Research into the shifts in microbial communities due to dietary changes or stress have led to discoveries in nutrition, behavior, and management of production animals (Chang et al., 2022; Diaz Heijtz et al., 2011; Fu et al., 2022). As an additional measure of bacterial communities, producers and researchers have increased focus on the role of beneficial bacteria on animal immune response and resilience of the community to perturbation by pathogenic microorganisms (Clavijo & Flórez, 2018; Desai et al., 2016; Haley et al., 2020; Mon et al., 2015).

Studies on the livestock microbiome may focus on characterizing the bacterial community with 16S rRNA gene amplicon sequencing or metagenomic analysis. Through analyzing the 16S rRNA genes of bacteria with gene amplicon sequencing or targeted metagenomic sequencing, researchers can identify the phylogenetic differences that categorize bacteria into various taxa and determine broad potential functions based on the taxonomic information. A study by Pitta et al. (2016) found differences in the rumen metagenome of dairy cattle during the stages of lactation, with late lactation having higher abundance of protein metabolism genes. Additionally, differences in carbohydrate enzymes found within the rumen metagenome were observed with a shift in the type of carbohydrate enzyme (e.g., debranching, cellulase, oligosaccharide degrading) between lactation periods (Pitta et al., 2016).

Studies on the bacterial communities of the sow reproductive tract emphasize the analysis of vaginal microbiota and the effects on sow production. Zhang et al. (2021) studied the vaginal microbiota in conjunction with gut microbiota to determine a connection between the microbial communities and a return to estrus. The authors found higher abundance of *Lactobacillus* in feces and *Ruminococcaceae* from vaginal swabs of sows with a normal return to estrus, two taxa associated with healthy communities for both environments, compared to sows that did not return

to estrus which had higher abundance of inflammatory and pathogenic bacterial taxa from vaginal swabs and feces (J. Zhang et al., 2021). In a study on differences in the vaginal microbiota of sows with or without urogenital infections, Poor et al. (2022) used culture-dependent (MALDI-TOF) and culture-independent (16S rRNA targeted metagenomics). Though there was similar characterization of genera between the culture-dependent and -independent methods, indicating a high abundance of the genera found within the samples, an identifiable difference in the genera between sows with and without urogenital infection was observed (Poor et al., 2022).

While analysis of bacterial communities focuses on identifying members and defining broad, overarching functions within higher order taxa, population analysis is concentrated on connecting function and phenotype within a specific genus, species, or strain. As discussed previously in this chapter, researchers have been able to isolate and characterize the biochemical properties of bacteria such as antibiotic resistance. Using the genome, the observable characteristics of bacteria can be linked to specific genetic information.

Traditionally, isolates were categorized by their phenotype. One of the earliest, and most famous, bacterial phenotype characterizations is the association of virulence and disease with colony morphology in *Streptococcus pneumoniae* by Frederick Griffiths in 1928 (Shen, 2019). Griffiths used classical microbiological techniques and determined what he called the "transforming principle" which is the ability for dead cells to transfer their phenotype to live cells. This set of observational experiments led the way for future scientists to understand the intricacies of DNA. Currently, we understand that the "smooth", virulent cells and "rough", avirulent cells have their morphology due to a set of genes that control the production of capsular polysaccharide – a common evasion mechanism against the host immune system shared among many virulent strains of bacteria (Klima et al., 2017; Sande & Whitfield, 2021). The theoretical knowledge of bacterial functions learned prior to the advent of high throughput sequencing still shapes how we approach experiments. However, the methods used reflect how much information has truly been gathered.

Following the growth and improvement of sequencing technologies, more bacteria have been characterized for phenotype and genotype, but there is still a large number of bacteria for which scientists do not have the ability to culture (Lloyd et al., 2018). Using the knowledge gained from mining bacterial genomes, proteomes, metagenomes, etc. from specific environments, the optimal conditions for culturing the "unculturable" can be ascertained but require deep datasets

that often do not exist (X. Wu et al., 2020). The connection of genotype and phenotype in well-known bacteria is ongoing as the bacteria evolve, but as more is learned about the complex interactions between members of a microbial community, increased emphasis is placed on characterizing the most influential members in conjunction with the most abundant.

Artificial intelligence-based analyses such as machine learning allows additional connection of genotype and phenotype. With machine learning, scientists are able to predict bacterial phenotypes using the genome and collected environmental information (Feldbauer, 2016; Lees et al., 2020). For example, the general function of rumen microbial communities is to ferment complex plant materials for further degradation by the host. However, the specific connection between which rumen bacteria are connected to rumen fermentation has recently been examined using machine learning. With the combination of sequence data and collected fermentation measurements from animals, J. Li et al. (2022) determined *Lactobacillus, Prevotellaceae, Selenomonas, Peptostreptococcus*, and *Olsenella* were significantly associated with gas production, nutrient digestibility, and volatile fatty acid production in sheep, goats, and cows.

1.7 Applying bacterial ecology to study communities and populations in livestock

However, there are many situations for which phenotype does not match current genotype information. Machine learning can be especially helpful in these situations, and the application of existing algorithms in bacterial ecology has pushed researchers to gather more information. One way machine learning has been used in cases of phenotype and genotype discordance is in the determination of antibiotic resistance. The pathogenic bacteria in animal diseases are often genera that have few representative genomes, resulting in unusable generalizations of current antibiotic resistance genotypes. Using machine learning, data collected on the specific animal pathogens can be used as the starting point for the discovery of antibiotic resistance determinants (Hicks et al., 2019; Lees et al., 2020).

While the different aspects of bacterial ecology are often studied individually, i.e., studies on populations or communities, the interactions between levels are all a part of a bigger picture. A review by Kodera et al. (2022) highlights the importance of understanding the interactions of bacteria within and outside of the target population. Though examining the individual potential of bacterial genera or species is helpful, it is important to remember that bacteria do not exist solely. The interactions between bacteria, biotic, and abiotic factors is the key focus of all ecological

studies, regardless of the level being examined (Mony et al., 2020; VanInsberghe et al., 2020). The utilization of traditional microbiology and ecology methods with current technology and analysis has improved our comprehension of many diseases in humans and will continue to be the way to advance understanding of the bacterial communities and populations within livestock.

The topics in this dissertation are split into two parts. In both parts of this dissertation, the information on bacteria is gathered with specific conditions in mind – bulls with satisfactory breeding soundness scores, farrowing piglets, and cattle with bovine respiratory disease. While some information can be generalized for similar conditions (e.g., the types of bacterial genera found within the respiratory tract), it is important to remember that ecological studies typically are observations at a specific point in time and/or under specific environmental conditions.

The first part of the dissertation is on the study of bacterial communities in the prepuce of bulls and the GIT of piglets during weaning. The objectives for these two studies are varied; examining bacterial communities of the reproductive tract bulls is to characterize healthy communities and how animal traits (age, breed, etc.) affect the types of bacteria that inhabit the reproductive organs. The objective of studying the GIT of piglets is to determine how supplementary dextrin can help ameliorate weaning stress.

The second part of the dissertation is the determination of antibiotic resistance using the genome of specific, pathogenic populations of bacteria from cattle with bovine respiratory disease. This research can be used for further investigation into how we can improve animal resilience to disease and overall health, as well as reduce the need for antibiotics, and improve. The objectives of the specific studies were to determine sequences predictive of phenotypic antibiotic resistance 1) from three BRD pathogens and 2) from M. haemolytica specifically.

1.8 References

Adnane, M., & Chapwanya, A. (2022). A review of the diversity of the genital tract microbiome and implications for fertility of cattle. *Animals*, *12*(4), 460. https://doi.org/10.3390/ani12040460

Albenberg, L. G., & Wu, G. D. (2014). Diet and the intestinal microbiome: associations, functions, and implications for health and disease. *Gastroenterology*, *146*(6), 1564–1572. https://doi.org/10.1053/j.gastro.2014.01.058

Allen, H. K., Levine, U. Y., Looft, T., Bandrick, M., & Casey, T. A. (2013). Treatment, promotion, commotion: Antibiotic alternatives in food-producing animals. *Trends in Microbiology*, *21*(3), 114–119. https://doi.org/10.1016/j.tim.2012.11.001

Anwar, H., Irfan, S., Hussain, G., Naeem Faisal, M., Muzaffar, H., Mustafa, I., Mukhtar, I., Malik, S., & Irfan Ullah, M. (2020). Gut microbiome: a new organ system in body. In G. Antonio Bastidas Pacheco & A. Ali Kamboh (Eds.), *Parasitology and Microbiology Research*. IntechOpen. https://doi.org/10.5772/intechopen.89634

Aruwa, C. E., Pillay, C., Nyaga, M. M., & Sabiu, S. (2021). Poultry gut health – microbiome functions, environmental impacts, microbiome engineering and advancements in characterization technologies. *Journal of Animal Science and Biotechnology*, *12*(1), 119. https://doi.org/10.1186/s40104-021-00640-9

Awad, W. A., Mann, E., Dzieciol, M., Hess, C., Schmitz-Esser, S., Wagner, M., & Hess, M. (2016). Age-related differences in the luminal and mucosa-associated gut microbiome of broiler chickens and shifts associated with *Campylobacter jejuni* infection. *Frontiers in Cellular and Infection Microbiology*, *6*. https://doi.org/10.3389/fcimb.2016.00154

Baker-Austin, C., Wright, M. S., Stepanauskas, R., & McArthur, J. V. (2006). Co-selection of antibiotic and metal resistance. *Trends in Microbiology*, *14*(4), 176–182. https://doi.org/10.1016/j.tim.2006.02.006

Bengtsson-Palme, J. (2018). The diversity of uncharacterized antibiotic resistance genes can be predicted from known gene variants—But not always. *Microbiome*, *6*(1), 125. https://doi.org/10.1186/s40168-018-0508-2

Bonten, M. J., Austin, D. J., & Lipsitch, M. (2001). Understanding the spread of antibiotic resistant pathogens in hospitals: Mathematical models as tools for control. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *33*(10), 1739–1746. https://doi.org/10.1086/323761

Bossi, L., Fuentes, J. A., Mora, G., & Figueroa-Bossi, N. (2003). Prophage contribution to bacterial population dynamics. *Journal of Bacteriology*, *185*(21), 6467–6471. https://doi.org/10.1128/JB.185.21.6467-6471.2003

Broom, L. J. (2017). The sub-inhibitory theory for antibiotic growth promoters. *Poultry Science*, *96*(9), 3104–3108. https://doi.org/10.3382/ps/pex114

Butaye, P., Devriese, L. A., & Haesebrouck, F. (2003). Antimicrobial growth promoters used in animal feed: effects of less well known antibiotics on Gram-positive bacteria. *Clinical Microbiology Reviews*, *16*(2), 175–188. https://doi.org/10.1128/CMR.16.2.175-188.2003

Cairns, J., Ruokolainen, L., Hultman, J., Tamminen, M., Virta, M., & Hiltunen, T. (2018). Ecology determines how low antibiotic concentration impacts community composition and horizontal transfer of resistance genes. *Communications Biology*, *1*(1). https://doi.org/10.1038/s42003-018-0041-7

Campbell, J. M., Crenshaw, J. D., & Polo, J. (2013). The biological stress of early weaned piglets. *Journal of Animal Science and Biotechnology*, *4*(1), 19. https://doi.org/10.1186/2049-1891-4-19

Card, R. M., Cawthraw, S. A., Nunez-Garcia, J., Ellis, R. J., Kay, G., Pallen, M. J., Woodward, M. J., & Anjum, M. F. (2017). An *in vitro* chicken gut model demonstrates transfer of a multidrug resistance plasmid from *Salmonella* to commensal *Escherichia coli*. *MBio*, *8*(4). https://doi.org/10.1128/mBio.00777-17

Carlson, D., Beattie, J. H., & Poulsen, H. D. (2007). Assessment of zinc and copper status in weaned piglets in relation to dietary zinc and copper supply. *Journal of Animal Physiology and Animal Nutrition*, *91*(1–2), 19–28. https://doi.org/10.1111/j.1439-0396.2006.00637.x

Chang, T., Ngo, J., Vargas, J. I., Rocheford, E., Rocheford, T., Ortiz, D., Karcher, D. M., & Johnson, T. A. (2022). Research note: orange corn altered the cecal microbiome in laying hens. *Poultry Science*, *101*(3), 101685. https://doi.org/10.1016/j.psj.2021.101685

Chen, H., Wang, C., Wang, Y., Chen, Y., Wan, M., Zhu, J., & Zhu, A. (2021). Effects of soft pellet creep feed on pre-weaning and post-weaning performance and intestinal development in piglets. *Animal Bioscience*, *34*(4), 714–723. https://doi.org/10.5713/ajas.20.0034

Choudhury, R., Middelkoop, A., Boekhorst, J., Gerrits, W. J. J., Kemp, B., Bolhuis, J. E., & Kleerebezem, M. (2021). Early life feeding accelerates gut microbiome maturation and suppresses acute post-weaning stress in piglets. *Environmental Microbiology*, *23*(11), 7201–7213. https://doi.org/10.1111/1462-2920.15791

Chuppava, B., Keller, B., Abd El-Wahab, A., Sürie, C., & Visscher, C. (2019). Resistance Reservoirs and multi-drug resistance of commensal *Escherichia coli* from excreta and manure isolated in broiler houses with different flooring designs. *Frontiers in Microbiology*, *10*, 2633. https://doi.org/10.3389/fmicb.2019.02633

Clavijo, V., & Flórez, M. J. V. (2018). The gastrointestinal microbiome and its association with the control of pathogens in broiler chicken production: A review. *Poultry Science*, *97*(3), 1006–1021. https://doi.org/10.3382/ps/pex359

Cox, G., & Wright, G. D. (2013). Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *International Journal of Medical Microbiology : IJMM*, *303*(6–7), 287–292. https://doi.org/10.1016/j.ijmm.2013.02.009

Cromwell, G. L. (2002). Why and how antibiotics are used in swine production. *Animal Biotechnology*, *13*(1), 7–27. https://doi.org/10.1081/ABIO-120005767

Daee, A. A., Khodakaram-Tafti, A., Derakhshandeh, A., & Seyedin, M. (2020). Identification of *Mycoplasma ovipneumoniae* and *Mycoplasma arginini* in sheep with pneumonia in North East of Iran. *Iranian Journal of Veterinary Research*, *21*(1), 15–19.

de Vries, M., van Middelaar, C. E., & de Boer, I. J. M. (2015). Comparing environmental impacts of beef production systems: A review of life cycle assessments. *Livestock Science*, *178*, 279–288. https://doi.org/10.1016/j.livsci.2015.06.020

Department of Health and Human Resources. (2015). Veterinary Feed Directive; Final Rule (Federal Register No. 106; Veterinary Feed Directive, pp. 31708–31735). Food and Drug Administration, HHS. https://www.govinfo.gov/content/pkg/FR-2015-06-03/pdf/2015-13393.pdf

Desai, M. S., Seekatz, A. M., Koropatkin, N. M., Kamada, N., Hickey, C. A., Wolter, M., Pudlo, N. A., Kitamoto, S., Terrapon, N., Muller, A., Young, V. B., Henrissat, B., Wilmes, P., Stappenbeck, T. S., Núñez, G., & Martens, E. C. (2016). A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell*, *167*(5), 1339-1353.e21. https://doi.org/10.1016/j.cell.2016.10.043

Dewi, G., & Kollanoor Johny, A. (2022). *Lactobacillus* in food animal production—A forerunner for clean label prospects in animal-derived products. *Frontiers in Sustainable Food Systems*, *6*, 831195. https://doi.org/10.3389/fsufs.2022.831195

Diaz Heijtz, R., Wang, S., Anuar, F., Qian, Y., Björkholm, B., Samuelsson, A., Hibberd, M. L., Forssberg, H., & Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3047–3052. https://doi.org/10.1073/pnas.1010529108

Dibner, J. J., & Richards, J. D. (2005). Antibiotic growth promoters in agriculture: History and mode of action. *Poultry Science*, *84*(4), 634–643. https://doi.org/10.1093/ps/84.4.634

Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A., & Crook, D. W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews Genetics*, *13*(9), 601–612. https://doi.org/10.1038/nrg3226

Dieterich, W., Schink, M., & Zopf, Y. (2018). Microbiota in the gastrointestinal tract. *Medical Sciences*, *6*(4), 116. https://doi.org/10.3390/medsci6040116

Donham, K. J. (2000). The concentration of swine production: effects on swine health, productivity, human health, and the environment. *Veterinary Clinics of North America: Food Animal Practice*, *16*(3), 559–597. https://doi.org/10.1016/S0749-0720(15)30087-6

Dowling, A., Dwyer, J. O., & Adley, C. C. (2017). Antibiotics: Mode of action and mechanisms of resistance. In *Antimicrobial research: Novel bioknowledge and educational programs* (pp. 536–545). Formatex Research Center.

Dudek, K., Nicholas, R. A. J., Szacawa, E., & Bednarek, D. (2020). *Mycoplasma bovis* infections—Occurrence, diagnosis and control. *Pathogens*, *9*(8), 640. https://doi.org/10.3390/pathogens9080640

Edrington, T. S., Harvey, R. B., Farrington, L. A., & Nisbet, D. J. (2001). Evaluation of subtherapeutic use of the antibiotics apramycin and carbadox on the prevalence of antimicrobial-resistant *Salmonella* infection in swine. *Journal of Food Protection*, *64*(12), 2067–2070. https://doi.org/10.4315/0362-028X-64.12.2067

Estrela, S., Sanchez-Gorostiaga, A., Vila, J. C., & Sanchez, A. (2021). Nutrient dominance governs the assembly of microbial communities in mixed nutrient environments. *ELife*, *10*, e65948. https://doi.org/10.7554/eLife.65948

Feldbauer, R. (2016). Machine learning for microbial phenotype prediction. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-14319-0

Food and Agriculture Organization of the United Nations. (2018). Future of food and agriculuture 2018: Alternative pathways to 2050. FOOD & AGRICULTURE ORG.

Food and Drug Administration. (2012). The judicious use of medically important antimicrobial drugs in food-producing animal (Guidance for Industry No. 209; Judicious Use of Antimicrobials). Food and Drug Administration. https://www.fda.gov/media/79140/download

Forum on Microbial Threats, Board on Global Health, Health and Medicine Division, & National Academies of Sciences, E. (2017). Combating antimicrobial resistance: A One Health approach to a global threat: Proceedings of a Workshop. https://doi.org/10.17226/24914

Fu, Y., Hu, J., Erasmus, M. A., Johnson, T. A., & Cheng, H. (2022). Effects of early-life cecal microbiota transplantation from divergently selected inbred chicken lines on growth, gut serotonin, and immune parameters in recipient chickens. *Poultry Science*, *101*(7), 101925. https://doi.org/10.1016/j.psj.2022.101925

Gerchman, I., Levisohn, S., Mikula, I., Manso-Silván, L., & Lysnyansky, I. (2011). Characterization of *in vivo*-acquired resistance to macrolides of *Mycoplasma gallisepticum* strains isolated from poultry. *Veterinary Research*, *42*(1), 90. https://doi.org/10.1186/1297-9716-42-90

Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature Medicine*, *24*(4), 392–400. https://doi.org/10.1038/nm.4517

Guo, C. Y., Ji, S. K., Yan, H., Wang, Y. J., Liu, J. J., Cao, Z. J., Yang, H. J., Zhang, W. J., & Li, S. L. (2020). Dynamic change of the gastrointestinal bacterial ecology in cows from birth to adulthood. *MicrobiologyOpen*, *9*(11). https://doi.org/10.1002/mbo3.1119

Haley, B. J., Kim, S.-W., Salaheen, S., Hovingh, E., & Van Kessel, J. A. S. (2020). Differences in the microbial community and resistome structures of feces from preweaned calves and lactating dairy cows in commercial dairy herds. *Foodborne Pathogens and Disease*, *17*(8), 494–503. https://doi.org/10.1089/fpd.2019.2768

Heeney, D. D., Gareau, M. G., & Marco, M. L. (2018). Intestinal *Lactobacillus* in health and disease, a driver or just along for the ride? *Current Opinion in Biotechnology*, *49*, 140–147. https://doi.org/10.1016/j.copbio.2017.08.004

Heinritz, S. N., Weiss, E., Eklund, M., Aumiller, T., Louis, S., Rings, A., Messner, S., Camarinha-Silva, A., Seifert, J., Bischoff, S. C., & Mosenthin, R. (2016). Intestinal microbiota and microbial metabolites are changed in a pig model fed a high-fat/low-fiber or a low-fat/high-fiber diet. *PLOS ONE*, *11*(4), e0154329-21. https://doi.org/10.1371/journal.pone.0154329

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: Surviving and thriving in the microbial jungle. *Nature Reviews Microbiology*, *8*(1), 15–25. https://doi.org/10.1038/nrmicro2259

Hicks, A. L., Wheeler, N., Sánchez-Busó, L., Rakeman, J. L., Harris, S. R., & Grad, Y. H. (2019). Evaluation of parameters affecting performance and reliability of machine learning-based antibiotic susceptibility testing from whole genome sequencing data. *PLOS Computational Biology*, *15*(9), e1007349. https://doi.org/10.1371/journal.pcbi.1007349

Hooper, L. V., Midtvedt, T., & Gordon, J. I. (2002). How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annual Review of Nutrition*, *22*, 283–307. https://doi.org/10.1146/annurev.nutr.22.011602.092259

Hsuan, S.-L., Liao, C.-M., Huang, C., Winton, J. R., Chen, Z.-W., Lee, W.-C., Liao, J.-W., Chen, T.-H., Chiou, C.-J., Yeh, K.-S., & Chien, M.-S. (2009). Efficacy of a novel *Pasteurella multocida* vaccine against progressive atrophic rhinitis of swine. *Vaccine*, *27*(22), 2923–2929. https://doi.org/10.1016/j.vaccine.2009.03.005

Hu, J., Chen, L., Zheng, W., Shi, M., Liu, L., Xie, C., Wang, X., Niu, Y., Hou, Q., Xu, X., Xu, B., Tang, Y., Zhou, S., Yan, Y., Yang, T., Ma, L., & Yan, X. (2018). *Lactobacillus frumenti* facilitates intestinal epithelial barrier function maintenance in early-weaned piglets. *Frontiers in Microbiology*, *9*, 897. https://doi.org/10.3389/fmicb.2018.00897

Jha, R., Fouhse, J. M., Tiwari, U. P., Li, L., & Willing, B. P. (2019). Dietary fiber and intestinal health of monogastric animals. *Frontiers in Veterinary Science*, *6*, 48. https://doi.org/10.3389/fvets.2019.00048

Johnston, A. M. (1998). Use of antimicrobial drugs in veterinary practice. *BMJ*, *317*(7159), 665–667. https://doi.org/10.1136/bmj.317.7159.665

Jutkina, J., Rutgersson, C., Flach, C.-F., & Larsson, D. G. J. (2016). An assay for determining minimal concentrations of antibiotics that drive horizontal transfer of resistance. *The Science of the Total Environment*, *548–549*, 131–138. https://doi.org/10.1016/j.scitotenv.2016.01.044

Kamath, P. L., Foster, J. T., Drees, K. P., Luikart, G., Quance, C., Anderson, N. J., Clarke, P. R., Cole, E. K., Drew, M. L., Edwards, W. H., Rhyan, J. C., Treanor, J. J., Wallen, R. L., White, P. J., Robbe-Austerman, S., & Cross, P. C. (2016). Genomics reveals historic and contemporary transmission dynamics of a bacterial disease among wildlife and livestock. *Nature Communications*, *7*(1), 11448. https://doi.org/10.1038/ncomms11448

Kay, K. L., Breidt, F., Fratamico, P. M., Baranzoni, G. M., Kim, G.-H., Grunden, A. M., & Oh, D.-H. (2017). *Escherichia coli* O157:H7 acid sensitivity correlates with flocculation phenotype during nutrient limitation. *Frontiers in Microbiology*, *8*, 1404. https://doi.org/10.3389/fmicb.2017.01404

Kim, M., Park, T., Jeong, J. Y., Baek, Y., & Lee, H.-J. (2020). Association between rumen microbiota and marbling score in Korean native beef cattle. *Animals: An Open Access Journal from MDPI*, *10*(4), E712. https://doi.org/10.3390/ani10040712

Klima, C. L., Zaheer, R., Briggs, R. E., & McAllister, T. A. (2017). A multiplex PCR assay for molecular capsular serotyping of *Mannheimia haemolytica* serotypes 1, 2, and 6. *Journal of Microbiological Methods*, *139*, 155–160. https://doi.org/10.1016/j.mimet.2017.05.010

Koch, F., Thom, U., Albrecht, E., Weikard, R., Nolte, W., Kuhla, B., & Kuehn, C. (2019). Heat stress directly impairs gut integrity and recruits distinct immune cell populations into the bovine intestine. *Proceedings of the National Academy of Sciences*, *116*(21), 10333–10338. https://doi.org/10.1073/pnas.1820130116

Kodera, S. M., Das, P., Gilbert, J. A., & Lutz, H. L. (2022). Conceptual strategies for characterizing interactions in microbial communities. *IScience*, *25*(2), 103775. https://doi.org/10.1016/j.isci.2022.103775

Kogut, M. H., & Arsenault, R. J. (2016). Editorial: Gut health: The new paradigm in food animal production. *Frontiers in Veterinary Science*, *3*. https://doi.org/10.3389/fvets.2016.00071

Kraimi, N., Dawkins, M., Gebhardt-Henrich, S. G., Velge, P., Rychlik, I., Volf, J., Creach, P., Smith, A., Colles, F., & Leterrier, C. (2019). Influence of the microbiota-gut-brain axis on behavior and welfare in farm animals: A review. *Physiology & Behavior*, *210*, 112658. https://doi.org/10.1016/j.physbeh.2019.112658

Lago, A., McGuirk, S. M., Bennett, T. B., Cook, N. B., & Nordlund, K. V. (2006). Calf respiratory disease and pen microenvironments in naturally ventilated calf barns in winter. *Journal of Dairy Science*, *89*(10), 4014–4025. https://doi.org/10.3168/jds.S0022-0302(06)72445-6

Leclercq, R., Cantón, R., Brown, D. F. J., Giske, C. G., Heisig, P., MacGowan, A. P., Mouton, J. W., Nordmann, P., Rodloff, A. C., Rossolini, G. M., Soussy, C. J., Steinbakk, M., Winstanley, T. G., & Kahlmeter, G. (2013). EUCAST expert rules in antimicrobial susceptibility testing. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, *19*(2), 141–160. https://doi.org/10.1111/j.1469-0691.2011.03703.x

Lees, J. A., Mai, T. T., Galardini, M., Wheeler, N. E., Horsfield, S. T., Parkhill, J., & Corander, J. (2020). Improved prediction of bacterial genotype-phenotype associations using interpretable pangenome-spanning regressions. *MBio*, *11*(4), e01344-20. https://doi.org/10.1128/mBio.01344-20

Li, J., Yan, H., Chen, J., Duan, C., Guo, Y., Liu, Y., Zhang, Y., & Ji, S. (2022). Correlation of ruminal fermentation parameters and rumen bacterial community by comparing those of the goat, sheep, and cow *in vitro*. *Fermentation*, *8*(9), 427. https://doi.org/10.3390/fermentation8090427

Li, Y., Wang, X., Wang, X., Wang, J., & Zhao, J. (2020). Life-long dynamics of the swine gut microbiome and their implications in probiotics development and food safety. *Gut Microbes*, *11*(6), 1824–1832. https://doi.org/10.1080/19490976.2020.1773748

Lim, J. Y., Yoon, J., & Hovde, C. J. (2010). A brief overview of *Escherichia coli* O157:H7 and its plasmid O157. *Journal of Microbiology and Biotechnology*, *20*(1), 5–14. https://doi.org/10.4014/jmb.0908.08007

Lima, T., Domingues, S., & Da Silva, G. J. (2020). Manure as a potential hotspot for antibiotic resistance dissemination by horizontal gene transfer events. *Veterinary Sciences*, *7*(3), 110. https://doi.org/10.3390/vetsci7030110

Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J., & Crosby, L. (2018). Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *MSystems*, *3*(5), e00055-18. https://doi.org/10.1128/mSystems.00055-18

Lopetuso, L. R., Scaldaferri, F., Petito, V., & Gasbarrini, A. (2013). Commensal *Clostridia*: Leading players in the maintenance of gut homeostasis. *Gut Pathogens*, *5*(1), 23. https://doi.org/10.1186/1757-4749-5-23

Mackenzie, D. D. S. (1967). Production and utilization of lactic acid by the ruminant. A Review. *Journal of Dairy Science*, *50*(11), 1772–1786. https://doi.org/10.3168/jds.S0022-0302(67)87715-4

Maes, D., Boyen, F., Haesebrouck, F., & Gautier-Bouchardon, A. V. (2020). Antimicrobial treatment of *Mycoplasma hyopneumoniae* infections. *The Veterinary Journal*, *259–260*, 105474. https://doi.org/10.1016/j.tvjl.2020.105474

Malmuthuge, N., Howell, A., Arsic, N., Prysliak, T., Perez-Casal, J., & Griebel, P. (2021). Effect of maternal separation and transportation stress on the bovine upper respiratory tract microbiome and the immune response to resident opportunistic pathogens. *Animal Microbiome*, *3*(1), 62. https://doi.org/10.1186/s42523-021-00123-2

Malone, A. S., Chung, Y.-K., & Yousef, A. E. (2006). Genes of *Escherichia coli* O157:H7 That are involved in high-pressure resistance. *Applied and Environmental Microbiology*, *72*(4), 2661–2671. https://doi.org/10.1128/AEM.72.4.2661-2671.2006

McLaren, M. R., & Callahan, B. J. (2020). Pathogen resistance may be the principal evolutionary advantage provided by the microbiome. *Philosophical Transactions of the Royal Society B: Biological Sciences, 375*(1808), 20190592. https://doi.org/10.1098/rstb.2019.0592

Mertens, D. R. (1997). Creating a system for meeting the fiber requirements of dairy cows. *Journal of Dairy Science, 80*(7), 1463–1481. https://doi.org/10.3168/jds.S0022-0302(97)76075-2

Middelkoop, A., Choudhury, R., Gerrits, W. J. J., Kemp, B., Kleerebezem, M., & Bolhuis, J. E. (2020). Effects of creep feed provision on behavior and performance of piglets around weaning. *Frontiers in Veterinary Science*, *7*, 520035. https://doi.org/10.3389/fvets.2020.520035

Mills, J. A. N., Crompton, L. A., Ellis, J. L., Dijkstra, J., Bannink, A., Hook, S., Benchaar, C., & France, J. (2014). A dynamic mechanistic model of lactic acid metabolism in the rumen. *Journal of Dairy Science, 97*(4), 2398–2414. https://doi.org/10.3168/jds.2013-7582

Mon, K. K. Z., Saelao, P., Halstead, M. M., Chanthavixay, G., Chang, H.-C., Garas, L., Maga, E. A., & Zhou, H. (2015). *Salmonella enterica* serovars Enteritidis infection alters the indigenous microbiota diversity in young layer chicks. *Frontiers in Veterinary Science*, *2*(2), 223. https://doi.org/10.3389/fvets.2015.00061

Monteiro, M., Poor, A., Muro, B., Carnevale, R., Leal, D., Garbossa, C., Moreno, A., & Almond, G. (2022). The sow microbiome: Current and future perspectives to maximize the productivity in swine herds. *Journal of Swine Health and Production*, *30*(4), 238–250. https://doi.org/10.54846/jshap/1277

Mony, C., Vandenkoornhuyse, P., Bohannan, B. J. M., Peay, K., & Leibold, M. A. (2020). A landscape of opportunities for microbial ecology research. *Frontiers in Microbiology*, *11*, 561427. https://doi.org/10.3389/fmicb.2020.561427

Mountzouris, K. C., Tsirtsikos, P., Kalamara, E., Nitsch, S., Schatzmayr, G., & Fegeros, K. (2007). Evaluation of the efficacy of a probiotic containing *Lactobacillus, Bifidobacterium, Enterococcus,* and *Pediococcus* strains in promoting broiler performance and modulating cecal microflora composition and metabolic activities. *Poultry Science, 86*(2), 309–317. https://doi.org/10.1093/ps/86.2.309

O'Hara, E., Neves, A. L. A., Song, Y., & Guan, L. L. (2020). The role of the gut microbiome in cattle production and health: driver or passenger? *Annual Review of Animal Biosciences*, *8*(1), 199–220. https://doi.org/10.1146/annurev-animal-021419-083952

Pacheco, A. R., Osborne, M. L., & Segrè, D. (2021). Non-additive microbial community responses to environmental complexity. *Nature Communications*, *12*(1), 2365. https://doi.org/10.1038/s41467-021-22426-3

Paone, P., & Cani, P. D. (2020). Mucus barrier, mucins and gut microbiota: The expected slimy partners? *Gut*, *69*(12), 2232–2243. https://doi.org/10.1136/gutjnl-2020-322260

Patra, A. K., & Aschenbach, J. R. (2018). Ureases in the gastrointestinal tracts of ruminant and monogastric animals and their implication in urea-N/ammonia metabolism: A review. *Journal of Advanced Research*, *13*, 39–50. https://doi.org/10.1016/j.jare.2018.02.005

Perez-Lopez, A., Behnsen, J., Nuccio, S.-P., & Raffatellu, M. (2016). Mucosal immunity to pathogenic intestinal bacteria. *Nature Reviews Immunology*, *16*(3), 135–148. https://doi.org/10.1038/nri.2015.17

Pitta, D. W., Indugu, N., Kumar, S., Vecchiarelli, B., Sinha, R., Baker, L. D., Bhukya, B., & Ferguson, J. D. (2016). Metagenomic assessment of the functional potential of the rumen microbiome in Holstein dairy cows. *Anaerobe*, *38*, 50–60. https://doi.org/10.1016/j.anaerobe.2015.12.003

Poor, A. P., Moreno, L. Z., Monteiro, M. S., Matajira, C. E. C., Dutra, M. C., Leal, D. F., Silva, A. P. S., Gomes, V. T. M., Barbosa, M. R. F., Sato, M. I. Z., & Moreno, A. M. (2022). Vaginal microbiota signatures in healthy and purulent vulvar discharge sows. *Scientific Reports*, *12*(1), 9106. https://doi.org/10.1038/s41598-022-13090-8

Quinteiro-Filho, W. M., Gomes, A. V. S., Pinheiro, M. L., Ribeiro, A., Ferraz-de-Paula, V., Astolfi-Ferreira, C. S., Ferreira, A. J. P., & Palermo-Neto, J. (2012). Heat stress impairs performance and induces intestinal inflammation in broiler chickens infected with *Salmonella* Enteritidis. *Avian Pathology*, *41*(5), 421–427. https://doi.org/10.1080/03079457.2012.709315

Ramos, B., Rosalino, L. M., Palmeira, J. D., Torres, R. T., & Cunha, M. V. (2022). Antimicrobial resistance in commensal *Staphylococcus aureus* from wild ungulates is driven by agricultural land cover and livestock farming. *Environmental Pollution*, *303*, 119116. https://doi.org/10.1016/j.envpol.2022.119116

Rangel, J. M., Sparling, P. H., Crowe, C., Griffin, P. M., & Swerdlow, D. L. (2005). Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. *Emerging Infectious Diseases*, *11*(4), 603–609. https://doi.org/10.3201/eid1104.040739

Rawat, N., Gilhare, V. R., Kushwaha, K. K., Hattimare, D. D., Khan, F. F., Shende, R. K., & Jolhe, D. K. (2019). Isolation and molecular characterization of *Mannheimia haemolytica* and *Pasteurella multocida* associated with pneumonia of goats in Chhattisgarh. *Veterinary World*, *12*(2), 331–336. https://doi.org/10.14202/vetworld.2019.331-336

Rosenfeld, C. S. (2017). Gut dysbiosis in animals due to environmental chemical exposures. *Frontiers in Cellular and Infection Microbiology*, *7*, 396. https://doi.org/10.3389/fcimb.2017.00396

Sande, C., & Whitfield, C. (2021). Capsules and extracellular polysaccharides in *Escherichia coli* and *Salmonella*. *EcoSal Plus*, *9*(2), eESP-0033-2020. https://doi.org/10.1128/ecosalplus.ESP-0033-2020

Sanglard, L. P., Schmitz-Esser, S., Gray, K. A., Linhares, D. C. L., Yeoman, C. J., Dekkers, J. C. M., Niederwerder, M. C., & Serão, N. V. L. (2020). Vaginal microbiota diverges in sows with low and high reproductive performance after porcine reproductive and respiratory syndrome vaccination. *Scientific Reports*, *10*(1), 3046. https://doi.org/10.1038/s41598-020-59955-8

Schabereiter-Gurtner, C., Piñar, G., Lubitz, W., & Rölleke, S. (2001). An advanced molecular strategy to identify bacterial communities on art objects. *Journal of Microbiological Methods*, *45*(2), 77–87. https://doi.org/10.1016/S0167-7012(01)00227-5

Segata, N., Boernigen, D., Tickle, T. L., Morgan, X. C., Garrett, W. S., & Huttenhower, C. (2013). Computational meta'omics for microbial community studies. *Molecular Systems Biology*, *9*(1), 666. https://doi.org/10.1038/msb.2013.22

Shabat, S. K. B., Sasson, G., Doron-Faigenboim, A., Durman, T., Yaacoby, S., Miller, M. E. B., White, B. A., Shterzer, N., & Mizrahi, I. (2016). Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *The ISME Journal*, *10*(12), 2958–2972. https://doi.org/10.1038/ismej.2016.62

Sharma, C., Rokana, N., Chandra, M., Singh, B. P., Gulhane, R. D., Gill, J. P. S., Ray, P., Puniya, A. K., & Panwar, H. (2018). Antimicrobial resistance: its surveillance, impact, and alternative management strategies in dairy animals. *Frontiers in Veterinary Science*, *4*, 237. https://doi.org/10.3389/fvets.2017.00237

Shen, C.-H. (2019). Nucleic acids: DNA and RNA. In *Diagnostic Molecular Biology* (pp. 1–25). Elsevier. https://doi.org/10.1016/B978-0-12-802823-0.00001-8

Sheppard, S. K., Guttman, D. S., & Fitzgerald, J. R. (2018). Population genomics of bacterial host adaptation. *Nature Reviews Genetics*, *19*(9), 549–565. https://doi.org/10.1038/s41576-018-0032-z

Shi, Z., Li, X., Wang, T., Xi, L., Cheng, P., Fang, M., & Liu, W. (2022). Application effects of three ventilation methods on swine in winter. *Agronomy Journal*, *114*(4), 1915–1922. https://doi.org/10.1002/agj2.20915

Shterzer, N., & Mizrahi, I. (2015). The animal gut as a melting pot for horizontal gene transfer. *Canadian Journal of Microbiology*, *61*(9), 603–605. https://doi.org/10.1139/cjm-2015-0049

Siqueira, F. M., Pérez-Wohlfeil, E., Carvalho, F. M., Trelles, O., Schrank, I. S., Vasconcelos, A. T. R., & Zaha, A. (2017). Microbiome overview in swine lungs. *PLOS ONE*, *12*(7), e0181503. https://doi.org/10.1371/journal.pone.0181503

Spadoni, I., Zagato, E., Bertocchi, A., Paolinelli, R., Hot, E., Di Sabatino, A., Caprioli, F., Bottiglieri, L., Oldani, A., Viale, G., Penna, G., Dejana, E., & Rescigno, M. (2015). A gut-vascular barrier controls the systemic dissemination of bacteria. *Science (New York, N.Y.)*, *350*(6262), 830–834. https://doi.org/10.1126/science.aad0135

Suchodolski, J. S. (2018). Gut brain axis and its microbiota regulation in mammals and birds. *Veterinary Clinics of North America: Exotic Animal Practice*, *21*(1), 159–167. https://doi.org/10.1016/j.cvex.2017.08.007

Summer, A., Lora, I., Formaggioni, P., & Gottardo, F. (2019). Impact of heat stress on milk and meat production. *Animal Frontiers*, *9*(1), 39–46. https://doi.org/10.1093/af/vfy026

Sundrum, A. (2015). Metabolic disorders in the transition period indicate that the dairy cows' ability to adapt is overstressed. *Animals*, *5*(4), 978–1020. https://doi.org/10.3390/ani5040395

Suresh, G., Das, R. K., Brar, S. K., Rouissi, T., Ramirez, A. A., Chorfi, Y., & Godbout, S. (2018). Alternatives to antibiotics in poultry feed: Molecular perspectives. *Critical Reviews in Microbiology*, *44*(3), 318–335. https://doi.org/10.1080/1040841X.2017.1373062

Tawfick, M. M., Elshamy, A. A., Mohamed, K. T., & El Menofy, N. G. (2022). Gut commensal *Escherichia coli*, a high-risk reservoir of transferable plasmid-mediated antimicrobial resistance traits. *Infection and Drug Resistance*, *Volume 15*, 1077–1091. https://doi.org/10.2147/IDR.S354884

Temple, D., & Manteca, X. (2020). Animal welfare in extensive production systems is still an area of concern. *Frontiers in Sustainable Food Systems*, *4*, 545902. https://doi.org/10.3389/fsufs.2020.545902

Tzortzis, G., Goulas, A. K., Gee, J. M., & Gibson, G. R. (2005). A novel galactooligosaccharide mixture increases the bifidobacterial population numbers in a continuous *in vitro* fermentation system and in the proximal colonic contents of pigs *in vivo*. *The Journal of Nutrition*, *135*(7), 1726–1731. https://doi.org/10.1093/jn/135.7.1726

Úbeda, C., Maiques, E., Knecht, E., Lasa, Í., Novick, R. P., & Penadés, J. R. (2005). Antibiotic-induced SOS response promotes horizontal dissemination of pathogenicity island-encoded virulence factors in staphylococci. *Molecular Microbiology*, *56*(3), 836–844. https://doi.org/10.1111/j.1365-2958.2005.04584.x

USDA - APHIS. (2002). Preventive practices in swine: administration of iron and antibiotics [Info Sheet]. https://www.aphis.usda.gov/animal_health/nahms/swine/downloads/swine2000/Swine2000_is_Iron.pdf

Valeriano, V. D. V., Balolong, M. P., & Kang, D.-K. (2017). Probiotic roles of *Lactobacillus* sp. in swine: Insights from gut microbiota. *Journal of Applied Microbiology*, *122*(3), 554–567. https://doi.org/10.1111/jam.13364

Van Cuong, N., Kiet, B. T., Phu, D. H., Van, N. T. B., Hien, V. B., Thwaites, G., Carrique-Mas, J., & Choisy, M. (2021). Effects of prophylactic and therapeutic antimicrobial uses in small-scale chicken flocks. *Zoonoses and Public Health*, *68*(5), 483–492. https://doi.org/10.1111/zph.12839

VanInsberghe, D., Arevalo, P., Chien, D., & Polz, M. F. (2020). How can microbial population genomics inform community ecology? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1798), 20190253. https://doi.org/10.1098/rstb.2019.0253

Villagómez-Estrada, S., Pérez, J. F., van Kuijk, S., Melo-Durán, D., Karimirad, R., & Solà-Oriol, D. (2020). Dietary preference of newly weaned pigs and nutrient interactions according to copper levels and sources with different solubility characteristics. *Animals*, *10*(7), 1133. https://doi.org/10.3390/ani10071133

Walter, J. (2008). Ecological role of lactobacilli in the gastrointestinal tract: Implications for fundamental and biomedical research. *Applied and Environmental Microbiology*, *74*(16), 4985–4996. https://doi.org/10.1128/AEM.00753-08

Wang, H., Xu, R., Zhang, H., Su, Y., & Zhu, W. (2020). Swine gut microbiota and its interaction with host nutrient metabolism. *Animal Nutrition*, *6*(4), 410–420. https://doi.org/10.1016/j.aninu.2020.10.002

Wang, J., Xiao, Y., Li, J., Qi, M., & Tan, B. (2021). Serum biochemical parameters and amino acids metabolism are altered in piglets by early-weaning and proline and putrescine supplementations. *Animal Nutrition*, *7*(2), 334–345. https://doi.org/10.1016/j.aninu.2020.11.007

Ward, M. J., Gibbons, C. L., McAdam, P. R., van Bunnik, B. A. D., Girvan, E. K., Edwards, G. F., Fitzgerald, J. R., & Woolhouse, M. E. J. (2014). Time-scaled evolutionary analysis of the transmission and antibiotic resistance dynamics of *Staphylococcus aureus* clonal complex 398. *Applied and Environmental Microbiology*, *80*(23), 7275–7282. https://doi.org/10.1128/AEM.01777-14

Wessels, A. G. (2022). Influence of the gut microbiome on feed intake of farm animals. *Microorganisms*, *10*(7), 1305. https://doi.org/10.3390/microorganisms10071305

Wickramasuriya, S. S., Park, I., Lee, K., Lee, Y., Kim, W. H., Nam, H., & Lillehoj, H. S. (2022). Role of physiology, immunity, microbiota, and infectious diseases in the gut health of poultry. *Vaccines*, *10*(2), 172. https://doi.org/10.3390/vaccines10020172

World Health Organization. (2019). Antimicrobial stewardship programmes in health-care facilities in low- and middle-income countries: A WHO practical toolkit. *JAC-Antimicrobial Resistance*, *1*(3), dlz072. https://doi.org/10.1093/jacamr/dlz072

Wu, J.-R., Shieh, H. K., Shien, J.-H., Gong, S.-R., & Chang, P.-C. (2003). Molecular characterization of plasmids with antimicrobial resistant genes in avian isolates of *Pasteurella multocida*. *Avian Diseases*, *47*(4), 1384–1392. https://doi.org/10.1637/z7035

Wu, X., Spencer, S., Gushgari-Doyle, S., Yee, M. O., Voriskova, J., Li, Y., Alm, E. J., & Chakraborty, R. (2020). Culturing of "unculturable" subsurface microbes: Natural organic carbon source fuels the growth of diverse and distinct bacteria from groundwater. *Frontiers in Microbiology*, *11*, 610001. https://doi.org/10.3389/fmicb.2020.610001

Wu, Y., Wang, Y., Yang, H., Li, Q., Gong, X., Zhang, G., & Zhu, K. (2021). Resident bacteria contribute to opportunistic infections of the respiratory tract. *PLOS Pathogens*, *17*(3), e1009436. https://doi.org/10.1371/journal.ppat.1009436

Xia, B., Wu, W., Fang, W., Wen, X., Xie, J., & Zhang, H. (2022). Heat stress-induced mucosal barrier dysfunction is potentially associated with gut microbiota dysbiosis in pigs. *Animal Nutrition*, *8*, 289–299. https://doi.org/10.1016/j.aninu.2021.05.012

Xu, J., Li, Y., Yang, Z., Li, C., Liang, H., Wu, Z., & Pu, W. (2018). Yeast probiotics shape the gut microbiome and improve the health of early-weaned piglets. *Frontiers in Microbiology*, *9*, 2011. https://doi.org/10.3389/fmicb.2018.02011

Yang, H., Xiang, Y., Robinson, K., Wang, J., Zhang, G., Zhao, J., & Xiao, Y. (2018). Gut microbiota is a major contributor to adiposity in pigs. *Frontiers in Microbiology*, *9*, 3045. https://doi.org/10.3389/fmicb.2018.03045

Yang, L., Liu, S., Ding, J., Dai, R., He, C., Xu, K., Honaker, C. F., Zhang, Y., Siegel, P., & Meng, H. (2017). Gut microbiota co-microevolution with selection for host humoral immunity. *Frontiers in Microbiology*, *8*, 1243. https://doi.org/10.3389/fmicb.2017.01243

Zhang, J., Liu, M., Ke, S., Huang, X., Fang, S., He, M., Fu, H., Chen, C., & Huang, L. (2021). Gut and vagina microbiota associated with estrus return of weaning sows and its correlation with the changes in serum metabolites. *Frontiers in Microbiology*, *12*, 690091. https://doi.org/10.3389/fmicb.2021.690091

Zhang, M., Chen, L., Ye, C., & Yu, X. (2017). Co-selection of antibiotic resistance via copper shock loading on bacteria from a drinking water bio-filter. *Environmental Pollution (Barking, Essex : 1987)*, *233*, 132–141. https://doi.org/10.1016/j.envpol.2017.09.084

PART I – BACTERIAL COMMUNITY ANALYSIS

CHAPTER 2. COMPOSITION AND DIVERSITY OF THE PREPUTIAL MICROBIOTA IN HEALTHY BULLS

Disclaimer: This chapter was previously published in the Journal of Theriogenology with the same title in March 2020.

Carmen L. Wickware[1], Timothy A. Johnson[1], and Jennifer H. Koziol[2,3]
[1]Deptartment of Animal Sciences and [2]Department of Veterinary Clinical Sciences, College of Veterinary Medicine, Purdue University, West Lafayette, IN
[3] This author has moved to Texas Tech University in Lubbock, Texas from the institution when the article was published

Acknowledgement: The Journal of Theriogenology provided permission for me to publish the above titled article. This study was funded by the Theriogenology Foundation, Montgomery, AL. We thank Ms. Rebecca Roloff for excellent technical assistance during analysis of the microbiome data.

Abstract

Characterization of microbial communities inhabiting the reproductive tracts of cattle may lead to a better comprehension of bovine physiology and reproductive health. To date, studies have been reported that have utilized culture-independent 16S ribosomal RNA (rRNA) for the classification of microbiota in the vaginal tract of cows but no studies have looked at the microbiota of the prepuce or penis of the bull. The aim of this study was to elucidate the microbiota present on the epithelial surface of the penis and prepuce of the post-pubertal bull using 16S rRNA gene sequencing. Ninety- two healthy bulls of a variety of ages and breeding history, presented for routine breeding soundness examinations, were utilized in this investigation. Bacteria belonging to *Firmicutes*, *Fusobacteria*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria* were identified in the prepuce. From all the bulls, two major community types were found, those with low or high bacterial species richness (up to 400 operational taxonomic units in one sample). There was no animal characteristic (breed or age) or management practice (feed type, antibiotic use, co-housing, breeding history) that was correlated with the bull penile microbial community composition. However, *Bradyrhizobium* is a distinguishing genus only found in the low diversity samples. The bull penile microbial community includes members of genera that are common in soil, cow vagina, respiratory tract, and feces. The baseline preputial microbial community in healthy bulls is

described in the current study. This knowledge can be used later when looking at how disease state alters the male reproductive microbial community.

2.1 Introduction

Worldwide there are approximately 971 million head of cattle with 92 million head within the United States alone. The production of cattle is one of the most important industries in the United States, accounting for $78.2 billion in cash receipts during 2015 (National Agriculutral Statistics Service, 2016). Despite the relevance of cattle economically, many aspects of their reproductive physiology and biology are still relatively unknown including the microbial community of certain body systems such as the urogenital tract of the male bovid. A microbial community can be defined as an assemblage of microorganisms present in a distinct environment. These microbial communities play a key role in maintaining health and altered microbial communities or dysbiosis have been associated with a variety of conditions (Nelson et al., 2010).

Up to this point, most microbial characterization of the bovine prepuce has involved investigation of urogenital disease. Bacterial genera isolated from the penis and prepuce and implicated in bovine reproductive disease including *Campylobacter*, *Chlamydia,* and *Histophilus* – specifically *H. somni* (Chaban et al., n.d.; Humphrey et al., n.d.; Parez, n.d.; Sandal & Inzana, 2010). These bacteria were isolated due to clinical signs of disease, but do not represent all of the species present in the bovine reproductive environment. In general, a majority of the microbial species within an environment resist cultivation in the laboratory (Amann et al., 1995; Bomar et al., 2011) and consequently have evaded notice during attempts to define microbial communities through bacterial culture. With the use of 16S ribosomal RNA (rRNA) amplification and sequencing, a larger majority of inhabiting microorganisms can be identified, allowing for the characterization of an entire microbial community of a selected environment. By identifying the symbiotic and dysbiotic states of the microbiota, the host-parasite-bacteria-virus complex interactions at mucosal or epithelial surfaces can be described allowing for important new research to be conducted.

The microbial community has been shown to have an impact on animal health and disease conditions as the host immune system is affected by changes in the microbiota and vice versa (Tomkovich & Jobin, 2016). Only recently has the vaginal and uterine microbiome been characterized in the cow, but the microbiota of the male urogenital system has yet to be determined

in any domestic animal species (Laguardia-Nascimento et al., 2015; Machado et al., 2012; Peng et al., 2013; Rodrigues et al., 2015; Swartz et al., 2014). In this study, samples from 92 healthy bulls were collected. This study defines the baseline microbial colonization of the bull prepuce and in the future may aid in determining factors governing disease conditions on the epithelial surface of the penis and prepuce. Similar work is beginning to be conducted in humans. For example, there is a clear distinction of urogenital microbial communities between males infected, or not, with sexually transmitted diseases or suffering from conditions such as prostatitis. Additionally, in humans the seminal microbial community correlates well with fertility in men (Weng et al., 2014). Through exploring and defining the bovid urogenital microbiota, the effects of the urogenital microbial community on host health and disease can be established.

2.2 Materials and Methods

2.2.1 Animals and Sample Collection

The objective of this study was to investigate and define the microbiota of the penis and prepuce of the post-pubertal bull using 16S rRNA profiling. Bulls presented to the Purdue University Food Animal Service for routine breeding soundness exams were utilized in this study pending client consent. This project was approved by the Purdue University IACUC committee (PRN-1701001529) and the Purdue University College of Veterinary Medicine Veterinary Clinical Studies Committee.

Ninety- two bulls were presented to the service between March and June 2017 and served as the source for the study group. All bulls were of breeding age with ages ranging between 15 months and nine years-of-age. A total of 7 breeds were represented: Angus (44), Crossbred (21) mainly Angus-Simmental crosses, Simmental (19), Hereford (5), Gelbvieh (1), Shorthorn (1), and Wagyu (1). A questionnaire was completed by the owner for each bull obtaining information on age, breed, diet, breeding history of bull co-housing, and antibiotic history from the previous year.

Bulls were restrained in a livestock squeeze chute per routine standards for performing routine breeding soundness examinations. The bull underwent a basic physical exam in-line with normal bull breeding soundness practices. During the process of semen collection by electroejaculation when the bull had fully extended his penis and prepuce a Dacron swab was used to collect samples from the epithelial surface of the penis and prepuce. Aseptic technique was used

to minimize contamination of the sample (i.e., gloves were worn during sampling, the sample was not grossly contaminated with dirt or debris, and once collected the sample was aseptically placed in a tube for transport). Once taken, the sample was quickly placed in individually labeled vials containing 1 mL RNAlater (Ambion, Austin, TX, USA) to stabilize and protect nucleic acids.

2.2.2 DNA Extraction and Sequencing

Samples were transported to the lab, and DNA was extracted from samples using the MagMAX CORE Nucleic Acid Purification Kit (Thermofisher Scientific Waltham, MA, USA) following the manufacturer's protocol. Extracted DNA was used for the construction of a 16S rRNA gene library following a standardized protocol (Kozich et al., 2013). Briefly, Illumina indexed reads were created using PCR amplification of the V4 region of bacterial 16S rRNA gene using the 515R (GTGCCAGCMGCCGCGGTAA) / 806R (GGACTACHVGGGTWTCTAAT) (Kozich et al., 2013). Amplification success was determined through gel electrophoresis as a quality check. No bands were observed in the negative control samples using water as the DNA template. Amplified DNA was normalized using a SequalPrep Normalization Plate (Invitrogen), and pooled into a single library. Library concentration was determined using the KAPA Library Quantification Kit (Roche) and library average fragment length was determined using the Bioanalyzer (Agilent) with a high sensitivity kit. Following the confirmation of proper DNA concentration, the pooled samples, mock community, and water, were sequenced (Illumina, MiSeq v2 kit, 500 cycle) at the Purdue Genomics Core Facility. PCR and sequencing quality was assessed by preparing 16S rRNA gene libraries for a known positive control mock community (20 Strain Even Mix Genomic Material; ATCC® MSA-1002TM) and water as a negative control (Kozich et al., 2013). Raw reads are available in the National Center for Biotechnology Information Sequence Read Archive (NCBI; SRA), Bioproject PRJNA534088, Biosamples SAMN11476796-891.

2.2.3 Sequence Analysis

Raw reads (1,175,143 read pairs from 92 bulls) were analyzed using mothur (v 1.39.3) (Kozich et al., 2013). The general pipeline for mothur is as follows: make contigs from raw reads, align contigs to reference sequences (SILVA database release 132) (Quast et al., 2012), screen and filter sequences to remove low quality reads (ambiguous bases allowed = 0, maximum read length

= 275, homopolymers allowed = 8), group sequences based on sequence similarity, classify sequences with reference to known taxonomic classifications (RDP training set 16) (Cole et al., 2013), cluster sequences, and run diversity metrics. After error removal processing, a total of 1,077,555 high quality sequences were obtained and used for downstream analysis (i.e., OTU picking and stats). Data was subsampled to 3000 reads. For purposes of reproducibility, all commands used in the mothur software are available at https://github.com/clwickwa/16S_Analysis.

2.2.4 Statistical Analysis

Because there was a lack of bulls aged 6 years and older (n=4, 1, and 1 for ages 6, 7, and 9, respectively) these were combined into one age group 6+; the same logic was used for low count breeds (Gelbvieh=1, Shorthorn=1, Hereford=4), which were combined into the breed category "Other". Alpha diversity, calculated by Chao1 or Shannon indices, was visualized with box-and-whisker plots using the R package "ggplot2" and alpha diversity differences were determined using ANOVA with a Tukey's test. Beta diversity was calculated using Bray-Curtis dissimilarity (Oksanen et al., 2018) and visualized with non-metric dimensional scaling (NMDS) plots using ggplot2. PERMANOVA was used to determine differences in beta diversity (age, breed, age x breed, diet, breeding history, and history of bull co-housing). Operational taxonomic unit (OTU) co-occurrence analysis was performed using OTU count data as input and calculation of all pairwise Spearman rank correlations between OTUs according to the methods described by Williams et al. (Williams et al., 2014). A multiple test correction q-value was calculated for alpha diversity Tukey's p-values and Spearman p-values using the Benjamini-Hochberg procedure to control the false discovery rate. All statistical analyses were done using R (R Core Team, 2013). To aid in computational reproducibility, all R scripts used in data analysis and visualization including alpha and beta diversity, ANOVA, and co-occurrence are available at https://github.com/clwickwa/16S_Analysis.

2.3 Results

2.3.1 α-Diversity: diversity within each sample

A total of 406 OTUs were identified. Rarefaction analysis indicated that the number of OTUs per sample reached a maximum, indicating that the sequencing effort captured the majority of samples' diversity (Figure 2.1). Estimated species richness (Chao1 index) in the bulls ranged from about 10 to 440 OTUs, though the majority of samples have 25-200 OTUs (Figure 2.2a). No statistically significant differences were seen in the microbial community due to age, diet or breed (Figure 2.2). β-Diversity: difference in community composition between samples

Two major clusters of samples were identified and the x-axis in the NMDS plot essentially divided these two groups of samples (Figure 2.3a): a tight cluster of samples on the left half of the plot and a disperse cluster of samples on the right half. Animal age, breed, diet, co-housing, or breeding history of the past year was not correlated with penile microbial community composition profile (Figure 2.3a). However, there was a significant difference in alpha diversity between these two clusters, with alpha diversity being significantly lower in the samples on the left (Figure 2.3b). Given that the alpha diversity is the most distinguishing factor of the sample clustering, samples with an axis 1 value less than zero on Figure 2.3a will be referred to as the low diversity (LD) samples and the samples with a positive axis 1 value on Figure 2.3a as the high diversity (HD) samples.

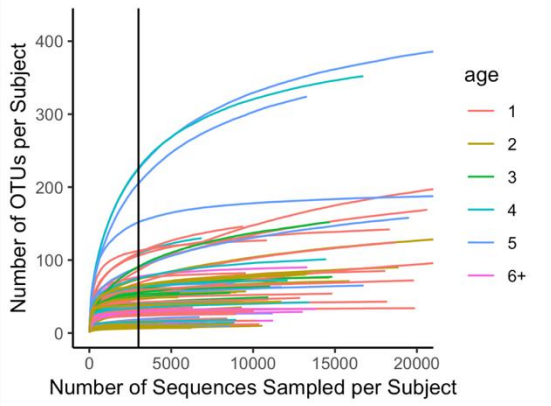

Figure 2.1 Rarefaction curves of observed bacterial species richness in bull preputial community by age

Species collection curves determined from sequence analysis. Each line represents one bull. Vertical line represents the subsampling cutoff: 3,000 sequences.

2.3.2 Taxonomic composition

The majority of 16S rRNA amplicons belonged to five main phyla including: *Firmicutes*, *Fusobacteria*, *Bacteroidetes*, *Proteobacteria*, and *Actinobacteria*. The genera found in the bull prepuce reveal a clear separation of bulls that had low overall alpha diversity, and samples that were dominated by more than two genera (e.g., *Fusobacterium*, *Histophilus*, *Porphyromonas,* and *Streptobacillus*) (Figure 2.4a). Difference in community membership was consistent with the clustering pattern observed in Figure 2.3a, as indicated, as well as the differences in alpha diversity. Bull samples from the HD cluster were evenly dominated by multiple genera (e.g.,*Fusobacterium*, *Histophilus*, *Porphyromonas,* and *Streptobacillus*) (Figure 2.4, right side).A wide spectrum of membership profiles were observed in the HD cluster. *Bradyrhizobium* was consistently present in samples from the low diversity cluster, but not the high diversity cluster (Figure 2.4b). Bull samples in the LD cluster showed lower abundance of minor members, while the membership profile in the HD cluster was not dominated by only one or two genera (Figure 2.4b).

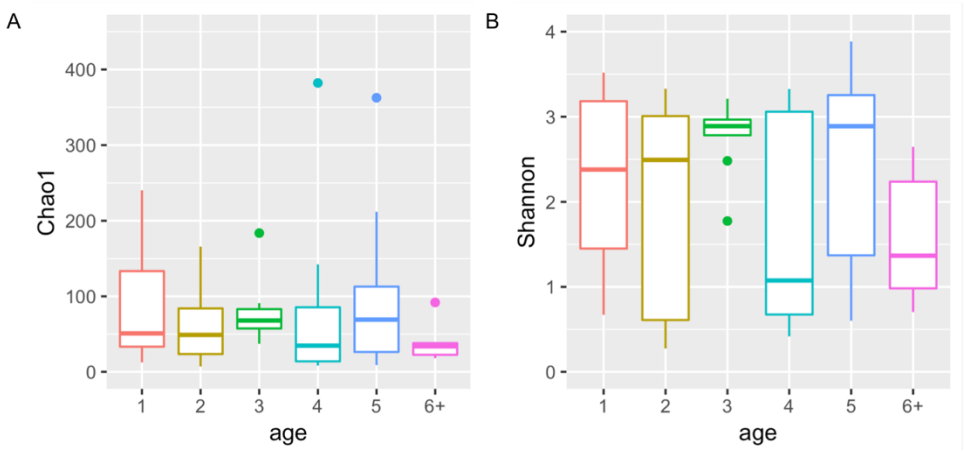

Figure 2.2 Alpha diversity of bull prepuce by age measured by Chao and Shannon indices

Box-and-whisker plots of diversity within samples. The lines of the box show the 25th, 50th, and 75th percentiles, the whiskers extend to the highest or lowest value within 1.5 times the interquartile range. Panel A: estimated number of OTUs (number of species found in samples) calculated by the Chao1 index. Panel B: estimated alpha diversity calculated by the Shannon index. No significant differences ($p > 0.05$) found. Age one n = 28, age two n = 19, age three n = 9, age four n = 12, age five n = 13, age 6+ n = 6.

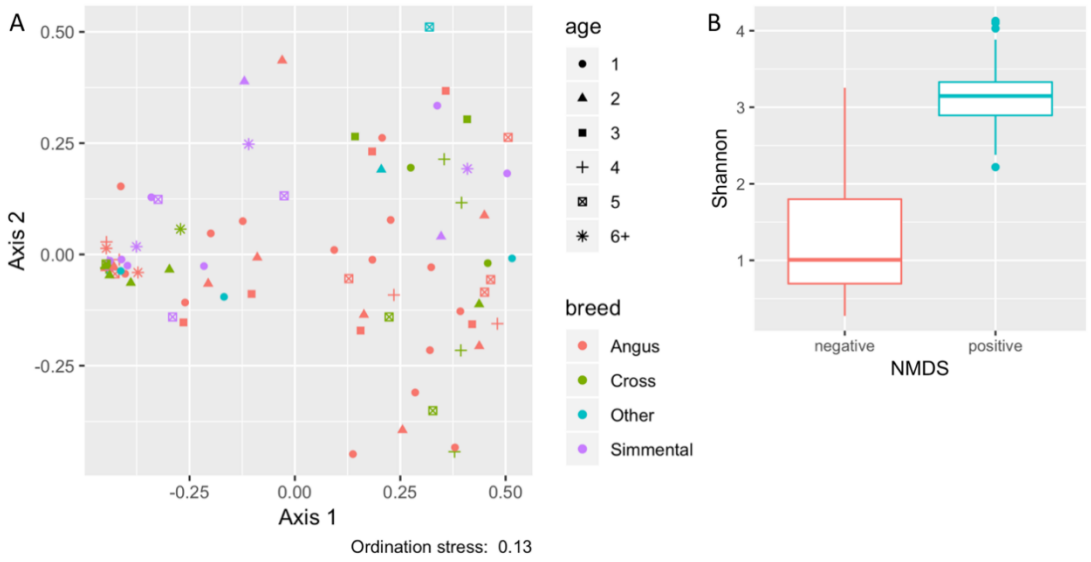

Figure 2.3. Diversity of bull preputial communities divided into low and high alpha diversity

NMDS ordination (A) showing beta diversity calculated with Bray-Curtis dissimilarity. Two clusters considered, samples with Axis 1 < 0.00 (negative) and samples with Axis 1 > 0.00 (positive). The alpha diversity of these two clusters (B) calculated with Shannon index was significant ($p < 0.05$).

2.3.3 Bacterial correlation

Strong co-occurrence patterns were observed between some members of the microbial communities. The strongest positive correlation was found between *Bacteroides* and unclassified *Clostridiales,* which are both high abundance OTUs from HD (Table 1). Additionally, *Bradyrhizobium*, mentioned previously as being consistently present in the low alpha diversity cluster, was negatively correlated (rho < -0.5, q < 0.05) with *Parvimonas* and unclassified *Ruminococcaceae*, genera found in the top-most abundant genera of the high diversity cluster (Apendix A.1). With the exception of the negative correlations between *Parvimonas* and unclassified *Ruminococcaceae* with *Bradyrhizobium*, all correlations found between OTUs were positive (Appendix A.1-A.8).

2.4 Discussion

In this study, the bull prepuce microbial community from healthy bulls was characterized in order to describe the baseline normal microbial community composition of the prepuce. These microbial communities do not appear to be shaped by diet, breed, age, farm or breeding history.

However, two community types were identified, one with low overall community diversity, and another type with higher community diversity.

Culture-dependant methods have identified many opportunistic bacterial pathogens colonizing the bull prepuce, and many of these populations were also detected in this study using culture-indepentant methods. For example, *Campylobacater fetus,* a bacterial species that causes bovine genital campylobacteriosis in females but has no known clinical signs in males, has been isolated from bull prepuce samples in other studies (Chaban et al., n.d.; Parez, n.d.) and *Campylobacter* was found in several of the high diversity samples (Figure 2.4b). Additionally, the genera *Histophilus, Porphyromonas,* and *Fusobacterium*, have all been previously isolated from cattle, and were present in high abundance in all samples with the exception of a few low diversity samples.

Histophilus was previously isolated from bull semen, urine, preputial washings as *Haemophilus somnus* (reclassified as *Histophilus somni*), while *Porphyromonas* and *Fusobacterium* have been isolated from non-prepuce, mucosal-associated environments in cattle, so it is not entirely surprising to find the genera in the prepuce (Elad et al., 2004; Humphrey et al., n.d.; Sandal & Inzana, 2010; Wolfe, 2018). These genera represent a small part of the microbial composition that was looked at in the current study. Thus, the culture-independent approach used in this study more comprehensively described the membership of the preputial microbial community than culture-dependent approaches have in the past.

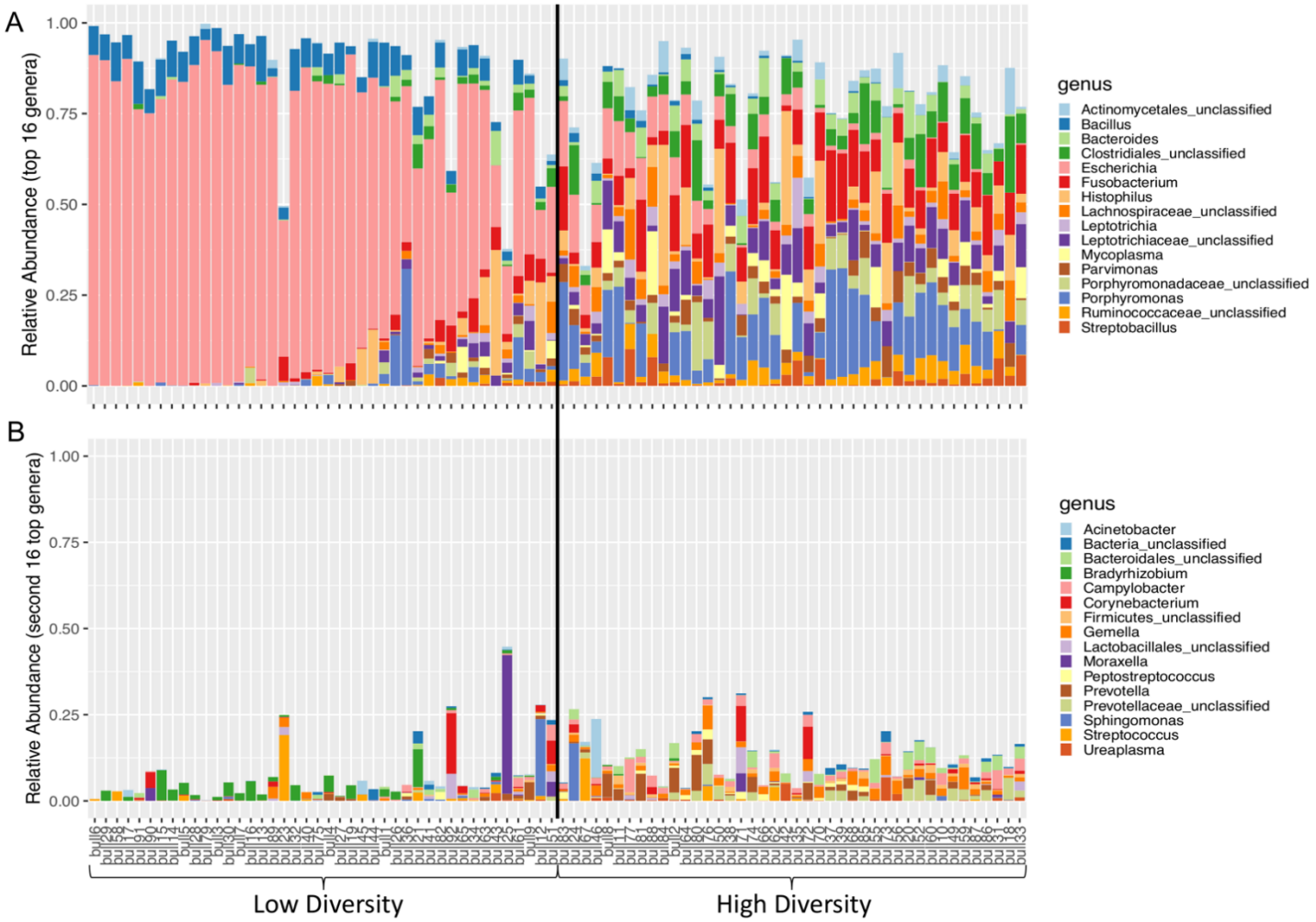

Figure 2.4. Relative abundance of genera in bull preputial bacterial community

Stacked bar graphs showing the relative abundance of genera (y-axis) for each bull sample (x-axis) sorted by the NMDS axis 1 with the vertical line representing 0.00 (Fig 3A). Panel A: The top 16 most abundant genera. Panel B: next 16 most abundant genera. Low Diversity = samples with NMDS axis 1 < 0.00, High Diversity = samples with NMDS axis 1 > 0.00.

Table 2.1 Significant Correlations of Bacterial Genera from Bul

OTU 1	OTU 2	pval	qval	rho	Genus 1	Genus 2
Otu0001	Otu0009	< 0.0001	< 0.0001	-0.78	Escherichia	Parvimonas
Otu0007	Otu0009	< 0.0001	< 0.0001	-0.76	Bacillus	Parvimonas
Otu0020	Otu0015	< 0.0001	< 0.0001	0.76	Ruminococcaceae_unclassified	Actinomycetales_unclassified
Otu0024	Otu0042	< 0.0001	< 0.0001	0.76	Fusobacterium	Firmicutes_unclassified
Otu0020	Otu0009	< 0.0001	< 0.0001	0.76	Ruminococcaceae_unclassified	Parvimonas
Otu0011	Otu0015	< 0.0001	< 0.0001	0.77	Lachnospiraceae_unclassified	Actinomycetales_unclassified
Otu0041	Otu0058	< 0.0001	< 0.0001	0.77	Corynebacterium	Lactobacillales_unclassified
Otu0019	Otu0042	< 0.0001	< 0.0001	0.77	Bacteroides	Firmicutes_unclassified
Otu0004	Otu0019	< 0.0001	< 0.0001	0.79	Porphyromonas	Bacteroides
Otu0012	Otu0042	< 0.0001	< 0.0001	0.80	Clostridiales_unclassified	Firmicutes_unclassified
Otu0019	Otu0012	< 0.0001	< 0.0001	0.83	Bacteroides	Clostridiales_unclassified
Otu0001	Otu0007	< 0.0001	< 0.0001	0.90	Escherichia	Bacillus
Otu0027	Otu0035	< 0.0001	< 0.0001	0.92	Porphyromonadaceae_unclassified	Porphyromonadaceae_unclassified

-This table shows the most significant correlations (rho > 0.5 or rho < - 0.5 and qval < 0.0001).
-OTU 1 corresponds to Genus 1 and OTU 2 corresponds to Genus 2.

2.4.1 Sources and selection of the preputial microbial community

In order to understand the composition of the bull prepuce microbial community, it is important to discuss its potential sources. It is assumed that the major sources of bacteria of the bull prepuce are soil, feces, cow vagina, and urine. The most distinguishing genus in the prepuce microbial community for the low diversity cluster was *Bradyrhizobium* (Figure 2.5). *Bradyrhizobium* are canonical soil bacteria that form symbiotic relationships with plant roots (Noisangiam et al., 2012) and has only been found to some significant amount in milk microbiome studies (Cremonesi et al., 2018; Taponen et al., 2019). Thus, soil is likely at least one source of bacteria to colonize the prepuce of bulls, especially of the LD cluster.

Some of the most distinguishing genera of the high diversity cluster include *Bacteroides*, unclassified *Ruminococcaceae*, *Histophilus,* and *Streptobacillus* (Figure 2.5). *Bacteroides* and *Ruminococcaceae* are commonly found in the bovine digestive tract (Yeoman et al., 2018). *Histophilus* has been isolated from urine as well as semen and prepuce (Humphrey et al., n.d.), and Swartz et al (Swartz et al., 2014) found that more than half of vaginal samples taken had *Streptobacillus,* perhaps the most common location for this genus in cattle. Thus soil, feces, urine, and the cow vagina may be potential sources of bacterial colonizers of the prepuce of bulls.

The selective conditions of the prepuce may have an even larger influence on the bacteria that colonize and grow on the prepuce than their environmental sources. A common characteristic of many of the genera identified is that they have been observed in mucosal environments. For instance, *Porphyromonas*, *Bacteroides*, and *Fusobacterium*, three of the most abundant genera in the HD bull samples, were among the most abundant genera from the upper respiratory tract, a mucosal environment, of healthy calves starting at three days after birth and are still present at 35 days (S. F. Lima et al., 2016). *Porphyromonas* was found colonizing the skin, oral cavity, upper GI, and the colon (Moe et al., 2010; K. Wang et al., 2016; Yano et al., 2010). *Mycoplasma*, *Leptotrichiaceae*, and *Fusobacterium* were among the five most abundant OTUs in lung and lymph node samples from cattle that died from bovine respiratory disease (D. Johnston et al., 2017). *Fusobacterium* spp. have been associated with human colorectal cancer and have be found in the placental microbial community, causing many cases of uterine disease in both humans and cattle (Kostic et al., 2013; Mor & Kwon, 2015). *Streptobacillus* was commonly found in the cow vagina (Swartz et al., 2014).

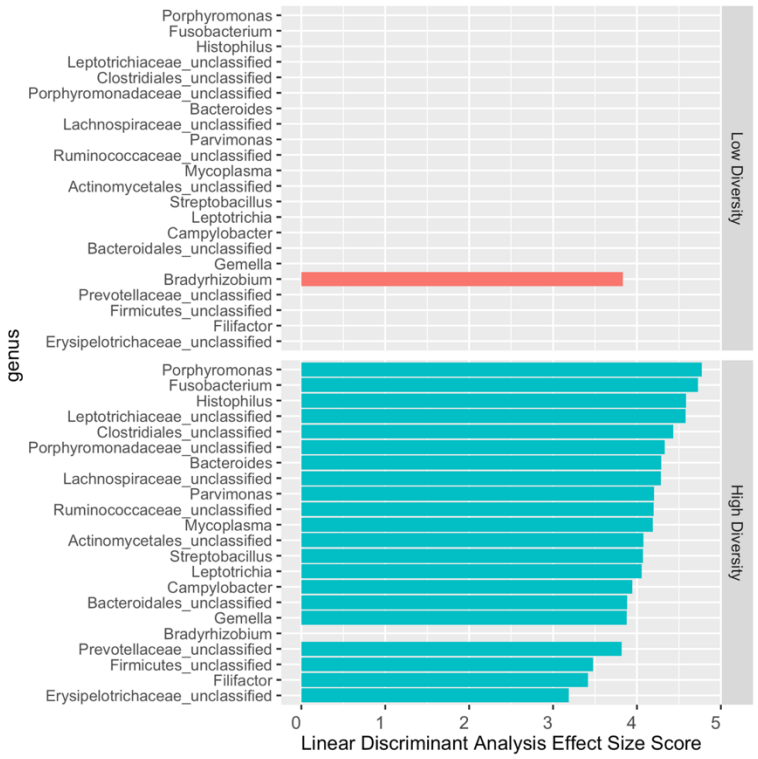

Figure 2.5. Linear Discriminant Analysis Effect Size of bulls with LD and HD bacterial communities

Bar plot showing the most distinguishing genera of either the low or high diversity cluster. Only OTUs with $p < 0.05$ and LDA > 2.0 were considered as being distinguishing.

Parvimonas is known to colonize the human oral cavity (Ang et al., 2013), but data on this genus in cattle is sparse. Like the prepuce, the environments discussed here – respiratory tract, oral cavity, placenta, and vagina – are also mucosal environments. While these mucosal environments are not necessarily physically connected, the environmental conditions may be the driving selective pressure for their colonization by similar organisms.

In additional to the mucosal surface as a selective pressure, the role of other factors in shaping the preputial bacterial community was investigated. Although no correlation was found between the bacteria and the microbial community composition and farm, age, diet, or breed of the bulls, two community types were identified. The presence of *Bradyrhizobium* appears to be a major driver of community composition, especially of the LD cluster. OTUs from *Bradyrhizobium* were negatively correlated with OTUs of high relative abundance in the HD samples. Thus, it appears that bacteria-bacteria interactions may play a role in the establishment of the prepuce microbiota. Future research could be performed to determine if *Bradyrhizobium* could be applied

to bulls as a probiotic to decrease the diversity of the microbiota on the bull prepuce to decrease the chance of colonization by opportunistic pathogens and possible parasitic disease. In addition to looking at the penile microbial community itself, the effect of low diversity versus high diversity penile microbiota on vaginal colonization and vaginal microbial community dysbiosis should be explored as the role of bacterial transfer in vaginal and uterine health pre- or post-breeding not yet known.

Bull preputial samples should continue to be collected and characterized using culture-independent methods to gain a more complete picture of the penile microbial community of healthy bulls, but it remains unclear if the penile microbial community is geographic-region specific. Additionally, the relationship between bacterial density and bacterial diversity should be exmined in the bull penile environment. The PCR negative control samples (water as DNA template) – which were initiated in the library preparation workflow after DNA extraction – had about 80% and 10% relative abundance of *Escherichia* and *Bacillus*, respectively (data not shown), similar in composition to the samples with the lowest alpha diversity (Figure 2.4a, far left). This observation makes it unclear if the community type is present in the bull prepuce or is an artifact of the DNA library preparation process. Despite the similarity of the negative control samples to the low diversity bull samples, the composition of the low diversity community type cannot be disregarded, based on the following observations: No PCR amplification products were observed following gel electrophoresis negative control PCR products, ii) PCR products resulting in weak observed electrophoresis bands were contained in both HD and LD groups and iii) other populations observed in the low diversity samples were not observed in the negative control samples, such as *Bradyrhizobium*. *Escherichia* and *Bacillus* are likely to be present in the bull penile microbial community. However, as a cautionary approach *Escherichia* and *Bacillus* were not extensively discussed as contributing factors to the composition and diversity of these samples. The use of sequencing negative control samples is helpful in determining possible contamination, but a lack of consensus remains as to how to incorporate data from negative controls with the biological samples (Kim et al., 2017). In future preputial microbial community studies, total bacterial load should be determined to understand if bacterial load and bacterial diversity are correlated.

Taken together, the taxonomic composition of the bull penile microbiome has similarities with other microbial communities within the bull. It appears that the source of the community

originates from soil, feces, and vaginal sources based on the shared occurrence of bacteria on the bull prepuce and these other environments. In addition, the similarity of the bull preputial microbial community with other mucosal microbial communities indicates that similar environmental conditions (moisture, nutrient contents, temperature) may exert similar selective pressures resulting in overlapping community compositions.

2.5 References

Amann, R. I., Ludwig, W., & Schleifer, K.-H. (1995). Phylogenetic identification and *In Situ* detection of individual microbial cells without cultivation. Microbiol. Rev., 59, 33.

Ang, M. Y., Dymock, D., Tan, J. L., Thong, M. H., Tan, Q. K., Wong, G. J., Paterson, I. C., & Choo, S. W. (2013). Genome sequence of *Parvimonas micra* strain A293, isolated from an abdominal abscess from a patient in the United Kingdom. Genome Announcements, 1(6). https://doi.org/10.1128/genomeA.01025-13

Bomar, L., Maltz, M., Colston, S., & Graf, J. (2011). Directed culturing of microorganisms using metatranscriptomics. MBio, 2(2), e00012-11. https://doi.org/10.1128/mBio.00012-11

Chaban, B., Chu, S., Hendrick, S., Waldner, C., & Hill, J. E. (n.d.). Evaluation of a *Campylobacter fetus* subspecies venerealis real-time quantitative polymerase chain reaction for direct analysis of bovine preputial samples. 8.

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., & Tiedje, J. M. (2013). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. Nucleic Acids Research, 42(D1), D633–D642. https://doi.org/10.1093/nar/gkt1244

Cremonesi, P., Ceccarani, C., Curone, G., Severgnini, M., Pollera, C., Bronzo, V., Riva, F., Addis, M. F., Filipe, J., Amadori, M., Trevisi, E., Vigo, D., Moroni, P., & Castiglioni, B. (2018). Milk microbiome diversity and bacterial group prevalence in a comparison between healthy Holstein, Friesian, and Rendena cows. PLOS ONE, 13(10), e0205054. https://doi.org/10.1371/journal.pone.0205054

Elad, D., Friedgut, O., Alpert, N., Stram, Y., Lahav, D., Tiomkin, D., Avramson, M., Grinberg, K., & Bernstein, M. (2004). Bovine necrotic vulvovaginitis associated with *Porphyromonas levii*. Emerging Infectious Diseases, 10(3), 505–507. https://doi.org/10.3201/eid1003.020592

Humphrey, J. D., Little, P. B., Barnum, D. A., Doig, P. A., Stephens, L. R., & Thorsen, J. (n.d.). Occurrence of "*Haemophilus somnus*" in bovine semen and in the prepuce of bulls and steers. Canadian Journal of Comprehensive Medicine, 3.

Johnston, D., Earley, B., Cormican, P., Murray, G., Kenny, D. A., Waters, S. M., McGee, M., Kelly, A. K., & McCabe, M. S. (2017). Illumina MiSeq 16S amplicon sequence analysis of bovine respiratory disease associated bacteria in lung and mediastinal lymph node tissue. BMC Veterinary Research, 13(1). https://doi.org/10.1186/s12917-017-1035-2

Kim, D., Hofstaedter, C. E., Zhao, C., Mattei, L., Tanes, C., Clarke, E., Lauder, A., Sherrill-Mix, S., Chehoud, C., Kelsen, J., Conrad, M., Collman, R. G., Baldassano, R., Bushman, F. D., & Bittinger, K. (2017). Optimizing methods and dodging pitfalls in microbiome research. 1–14. https://doi.org/10.1186/s40168-017-0267-5

Kostic, A. D., Chun, E., Robertson, L., Glickman, J. N., Gallini, C. A., Michaud, M., Clancy, T. E., Chung, D. C., Lochhead, P., Hold, G. L., El-Omar, E. M., Brenner, D., Fuchs, C. S., Meyerson, M., & Garrett, W. S. (2013). *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host & Microbe, 14(2), 207–215. https://doi.org/10.1016/j.chom.2013.07.007

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

Laguardia-Nascimento, M., Branco, K. M. G. R., Gasparini, M. R., Giannattasio-Ferraz, S., Leite, L. R., Araujo, F. M. G., Salim, A. C. de M., Nicoli, J. R., de Oliveira, G. C., & Barbosa-Stancioli, E. F. (2015). Vaginal microbiome characterization of nellore cattle using metagenomic analysis. PLOS ONE, 10(11), e0143294. https://doi.org/10.1371/journal.pone.0143294

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). https://doi.org/10.1038/srep29050

Machado, V. S., Oikonomou, G., Bicalho, M. L. S., Knauer, W. A., Gilbert, R., & Bicalho, R. C. (2012). Investigation of postpartum dairy cows' uterine microbial diversity using metagenomic pyrosequencing of the 16S rRNA gene. Veterinary Microbiology, 159(3–4), 460–469. https://doi.org/10.1016/j.vetmic.2012.04.033

Moe, K. K., Yano, T., Misumi, K., Kubota, C., Nibe, K., Yamazaki, W., Muguruma, M., & Misawa, N. (2010). Detection of antibodies against *Fusobacterium necrophorum* and *Porphyromonas levii*-like species in dairy cattle with papillomatous digital dermatitis. Microbiology and Immunology, 54(6), 338–346. https://doi.org/10.1111/j.1348-0421.2010.00220.x

Mor, G., & Kwon, J.-Y. (2015). Trophoblast-microbiome interaction: A new paradigm on immune regulation. American Journal of Obstetrics and Gynecology, 213(4), S131–S137. https://doi.org/10.1016/j.ajog.2015.06.039

National Agriculutral Statistics Service. (2016). Overview of the United States cattle industry. USDA. https://downloads.usda.library.cornell.edu/usda-esmis/files/h702q636h/1n79h639b/1z40kw116/Catt-01-29-2016.pdf

Nelson, D. E., Van Der Pol, B., Dong, Q., Revanna, K. V., Fan, B., Easwaran, S., Sodergren, E., Weinstock, G. M., Diao, L., & Fortenberry, J. D. (2010). Characteristic male urine microbiomes associate with asymptomatic sexually transmitted infection. PLoS ONE, 5(11), e14116. https://doi.org/10.1371/journal.pone.0014116

Noisangiam, R., Teamtisong, K., Tittabutr, P., Boonkerd, N., Toshiki, U., Minamisawa, K., & Teaumroong, N. (2012). Genetic diversity, symbiotic evolution, and proposed infection process of *Bradyrhizobium* strains isolated from root nodules of *Aeschynomene americana* L. in Thailand. Applied and Environmental Microbiology, 78(17), 6236–6250. https://doi.org/10.1128/AEM.00897-12

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., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2018). vegan: community ecology package (2.5-2) [Computer software]. https://CRAN.R-project.org/package=vegan

Parez, M. (1985). The most important genital diseases of cattle (control, treatment, and the hygiene of semen collection). 19.

Peng, Y., Wang, Y., Hang, S., & Zhu, W. (2013). Microbial diversity in uterus of healthy and metritic postpartum Holstein dairy cows. Folia Microbiologica, 58(6), 593–600. https://doi.org/10.1007/s12223-013-0238-6

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glockner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/

Rodrigues, N. F., Kästle, J., Coutinho, T. J. D., Amorim, A. T., Campos, G. B., Santos, V. M., Marques, L. M., Timenetsky, J., & de Farias, S. T. (2015). Qualitative analysis of the vaginal microbiota of healthy cattle and cattle with genital-tract disease. Genetics and Molecular Research, 14(2), 6518–6528. https://doi.org/10.4238/2015.June.12.4

Sandal, I., & Inzana, T. J. (2010). A genomic window into the virulence of *Histophilus somni*. Trends in Microbiology, 18(2), 90–99. https://doi.org/10.1016/j.tim.2009.11.006

Swartz, J. D., Lachman, M., Westveer, K., ONeill, T., Geary, T., Kott, R. W., Berardinelli, J. G., Hatfield, P. G., Thomson, J. M., Roberts, A., & Yeoman, C. J. (2014). Characterization of the vaginal microbiota of ewes and cows reveals a unique microbiota with low levels of Lactobacilli and near-neutral ph. Frontiers in Veterinary Science, 1. https://doi.org/10.3389/fvets.2014.00019

Taponen, S., McGuinness, D., Hiitiö, H., Simojoki, H., Zadoks, R., & Pyörälä, S. (2019). Bovine milk microbiome: A more complex issue than expected. Veterinary Research, 50(1), 44. https://doi.org/10.1186/s13567-019-0662-y

Tomkovich, S., & Jobin, C. (2016). Microbiota and host immune responses: A love-hate relationship. Immunology, 147(1), 1–10. https://doi.org/10.1111/imm.12538

Wang, K., Lu, W., Tu, Q., Ge, Y., He, J., Zhou, Y., Gou, Y., Van Nostrand, J. D., Qin, Y., Li, J., Zhou, J., Li, Y., Xiao, L., & Zhou, X. (2016). Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus. Scientific Reports, 6(1). https://doi.org/10.1038/srep22943

Weng, S.-L., Chiu, C.-M., Lin, F.-M., Huang, W.-C., Liang, C., Yang, T., Yang, T.-L., Liu, C.-Y., Wu, W.-Y., Chang, Y.-A., Chang, T.-H., & Huang, H.-D. (2014). Bacterial communities in semen from men of infertile couples: Metagenomic sequencing reveals relationships of seminal microbiota to semen quality. PLoS ONE, 9(10), e110152. https://doi.org/10.1371/journal.pone.0110152

Williams, R. J., Howe, A., & Hofmockel, K. S. (2014). Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. Frontiers in Microbiology, 5. https://doi.org/10.3389/fmicb.2014.00358

Wolfe, D. F. (2018). Review: Abnormalities of the bull – occurrence, diagnosis and treatment of abnormalities of the bull, including structural soundness. Animal, 12(s1), s148–s157. https://doi.org/10.1017/S1751731118000939

Yano, T., Moe, K. K., Yamazaki, K., Ooka, T., Hayashi, T., & Misawa, N. (2010). Identification of candidate pathogens of papillomatous digital dermatitis in dairy cattle from quantitative 16S rRNA clonal analysis. Veterinary Microbiology, 143(2–4), 352–362. https://doi.org/10.1016/j.vetmic.2009.12.009

Yeoman, C. J., Ishaq, S. L., Bichi, E., Olivo, S. K., Lowe, J., & Aldridge, B. M. (2018). Biogeographical differences in the influence of maternal microbial sources on the early successional development of the bovine neonatal gastrointestinal tract. Scientific Reports, 8(1), 3197. https://doi.org/10.1038/s41598-018-21440-8

CHAPTER 3. EFFECTS OF PRE- AND POST-WEANING SOLUBLE FIBER (DEXTRIN) SUPPLEMENTATION ON THE INTESTINAL MICROBIAL COMMUNITY OF PIGS

Carmen L. Wickware*[†], Clayton S. Chastain*[†], Allan P. Schinckel*, Brian T. Richert*, J. Scott Radcliffe*[,1], and Timothy A. Johnson*[,1]

*Department of Animal Sciences, Purdue University, 270 S. Russell St., West Lafayette, IN 47907
[†]Contributed equally; Clayton S. Chastain responsible for sample collection and analysis for performance of feed efficiency, histology, and short chain fatty acid analysis; Carmen L. Wickware responsible for sample collection and analysis for performance of microbial community analysis, and manuscript preparation.

Abstract

Digestive dysbiosis, impaired gut barrier function, and disease are common following weaning in swine. Post-weaning, the gut microbial community is significantly altered, partially due to the change in diet. The goal of this study was to determine the effects of supplemental soluble fiber (dextrin) administered both pre- and post-weaning on growth performance, short chain fatty acid (SCFA) production, and bacterial community composition. Pigs (n = 40) were blocked by genetics and body weight (BW), and randomly allotted to treatments in a 2 x 2 factorial design with or without supplemental dextrin pre-weaning and with or without supplemental dextrin post-weaning in a 35-day (d) experiment. Dextrin was suspended in chocolate milk and administered orally from 14 d prior to weaning until 4 d post-weaning, after which it was added in the diet at 1% by weight. Growth performance was not affected ($P > 0.10$) by treatment. When considering SCFAs, pigs fed fiber only in the nursery had the highest fecal acetate concentration, while pigs never receiving fiber had the lowest concentration. The intestinal microbiome was altered due to dextrin treatment ($P < 0.001$) and microbiota shifts supported the SCFA analysis; pigs that received fiber pre- and/or post-weaning had an increase in some OTUs that belong to genera known to produce SCFAs, such as *Coprococcus*, *Butyricimonas*, *Mogibacterium*, and *Paraprevotella*, as well as other beneficial commensal bacteria. This study shows the importance of fiber in shaping the gut health of swine, especially during the pre-weaning period, and merits continued study to determine the optimal fibers that should be used, the method of fiber supplementation.

Keywords: fiber, gastrointestinal, microbiome, mucosal, swine, weaning stress

LIST OF ABBREVIATIONS

ADFI - average daily feed intake

ADG - average daily gain

BW - body weight

G:F - gain:feed ratio

GI - gastrointestinal

LAB - lactic acid bacteria

PLI - proximal large intestine

q value - false discovery rate corrected p value

SCFA - short chain fatty acids

3.1 Introduction

Weaning is one of the most stressful periods in the life of a pig and is associated with many gastrointestinal (GI) alterations including a reduced villus height to crypt depth ratio (villus:crypt), changes to the intestinal microbiome, and poor regulation of an underdeveloped immune system (Pluske et al., 1997; Heo et al., 2013). These alterations to the pig GI tract often cause diarrhea, decreased absorptive capacity, increased intestinal inflammation, and decreased growth performance following weaning (Campbell et al., 2013; Pluske et al., 1996; Wijtten et al., 2011), likely due to the sudden environmental, social, and dietary changes that occur around the time of weaning (Hötzel et al., 2011; Lallès et al., 2007), which can then lead to animal stress (Moeser et al., 2007) and decreased feed intake (Pluske et al., 1997). Higher incidence of diarrhea following weaning have been linked to the proliferation of enterotoxigenic or enteropathogenic *Escherichia coli* (Rhouma et al., 2017), which may be exacerbated by decreased barrier function causing an imbalanced microbiome and pathogenic infection (H. Chen et al., 2013).

Intestinal barrier function is critical to animal health especially at weaning. Intestinal permeability can be increased due to reduced tight junction protein (e.g., occludin) expression or concentration in the intestinal mucosa (Oswald, 2006), which can allow bacteria and other toxic substances to pass through the intestinal barrier paracellularly, activating the immune system (Turner, 2006), leading to intestinal inflammation (Campbell et al., 2013). Intestinal permeability and cellular integrity can also be altered by pro-inflammatory cytokines (McKay and Baird, 1999). The intestinal microbiota are also thought to be important to intestinal barrier function and the overall health and development of the pig (Frese et al., 2015a). Short chain fatty acids (SCFAs) are produced by beneficial bacteria, and an alteration in SCFA concentrations typically indicates a shift in the bacterial community composition (Franklin et al., 2002).

Antibiotics are often added to the feed post-weaning to prevent disease. However, antibiotic use in livestock production is coming under increased public scrutiny due to concerns over antibiotic resistance, for both humans and pigs. For these reasons, there has been a push by the animal industry to reduce antibiotic usage, mostly to satisfy market demand for meat produced without antibiotics (Singer et al., 2019). To reduce the need for antibiotics, a multitude of new strategies have been researched to decrease intestinal permeability and diarrhea while increasing absorptive capacity and maintaining growth performance achieved similar to when animals are

treated with antibiotics (Allen et al., 2013b), but many challenges remain in this new agricultural system and no convincing alternatives to antibiotics have been found (Allen, 2017).

The microbiome of swine undergoes a significant shift at the time of weaning due to the dietary shift from a milk-based diet to a cereal grain-based diet. Significant shifts in alpha and beta diversity have been observed post-weaning in piglets (Frese et al., 2015a; Guevarra et al., 2018). The bacterial families *Ruminococcaceae, Prevotellaceae,* and *Lactobacillaceae* increase post weaning in the healthy pig (Frese et al., 2015a) in response to a plant-based (including fibers) diet post weaning. In humans, inclusion of dietary fiber can alter the gut microbiota composition, acting as a prebiotic energy source (Bauer et al., 2006). In pigs, fiber can improve gut health and prevent diarrhea in growing pigs (Jha et al., 2019), but is usually only given to gestating sows to control caloric intake, enhance gut fill, and/or reduce feed cost (Jarrett & Ashworth, 2018). Little is known about the impact of supplemental fiber on the suckling piglet. Creep feed is thought to have beneficial effects and help train piglets to explore and eat solid food. However, current studies have conflicting evidence of the effects of creep feed on animal performance with some suggesting there is no improvement (Christensen & Huber, 2021; Middelkoop et al., 2020), while others showed a benefit in animal performance (H. Chen et al., 2021; Lee et al., 2021). Regardless, access to solid feed prior to weaning can alter the gut microbiome composition (Choudhury et al., 2021; Shim et al., 2005), thus we hypothesized that supplementation of dextrin in the diet of the pre-weaned piglet would result in an enhanced gut microbiome at weaning, and thereby reduce weaning stress to the gut.

Soluble fibers, specifically, have been reported to reduce lesions caused by inflammation in both the ileum and colon, influence expression of anti-inflammatory cytokines in mice (Bassaganya-Riera et al., 2011), decrease fecal pH and increase calcium absorption (Whisner et al., 2016). Five soluble fibers were shown to increase the abundance of *Bifidobacteria* and *Lactobacillus* species in vitro, in which fermentation of dextrin was found to generate the most SCFAs compared with the other four soluble fibers (Maathuis et al., 2009). In rats, dextrin was reported to benefit gut health by increasing cecal weight and goblet cell count, and alter mucin composition (Knapp et al., 2013). It is thought that these beneficial effects are brought about by altering the gut microbial community to a more beneficial community that produces more SCFAs. The purpose of this study was to supplement the diet of pigs pre- and/or post-weaning with dextrin

to test the ability of a soluble fiber to modify the gut microbiome composition and/or alleviate negative health symptoms that occur shortly after weaning.

3.2 Materials and methods

3.2.1 Animals and experimental design

All animal procedures were approved by the Purdue Animal Care and Use Committee (PACUC #1303000841). A total of 40 barrows (Landrace x Chesterwhite cross (Calbrix et al., 2012)) were blocked by BW and genetics, and randomly allotted to a 2 x 2 factorial experiment with or without supplemental dextrin fiber pre-weaning and with or without supplemental dextrin fiber post-weaning. At the beginning of the study, pigs were 9.2 ± 1 d of age. The study lasted for 5 weeks: beginning 14 d prior to weaning (d -14) and ending 21 d post-weaning (d 21). From d -14 to d 0, pigs were housed with the sow in their respective farrowing crate. On d 0, pigs were weaned and moved to group housed pens (8 pigs / pen), and on d 4 pigs were moved to individual pens where they remained for the remainder of the experiment.

Dextrin (Equate, Bentonville, AR) was given to pigs orally (dissolved in chocolate milk) at a rate of 1 g/d on days -14 to -8, 2 g/d on d -7 to -1 and 3 g/d on d -1 until d 4 relative to weaning (d 0) to mimic the consumption of creep feed (i.e., the concentration of dextrin was increased as the animals matured from farrowing to nursery). Chocolate milk was used to aid in animal acceptance of swallowing. The target fiber dose was 0.05 g dextrin / kg BW / d according to doses in human trials (Whisner et al., 2016). Starting on day 4, dextrin was mixed into the feed at 1% of the diet. Pigs not receiving fiber during the oral dosing period received chocolate milk with no added fiber to mimic the same level of stress in all animals.

After weaning on d 0, all pigs were fed a common Phase 1 nursery diet with no supplemental dextrin added. During this time (d 0-4), pigs on dextrin treatments were given the fiber orally, dissolved in chocolate milk. All diets were formulated to meet or exceed all nutrient requirements (Table 3.1) according to the NRC (2012). Diets were analyzed at Purdue University for energy, DM, ash, nitrogen, phosphorous, NDF, and ADF (Table 3.2). The Phase 1 diet (corn-soy based diet with lactose and animal protein sources) had no supplemental dextrin (Table 3.1).

Table 3.1 Formulated composition of swine diets

	Phase 1		Phase 2		Phase 3	
Ingredient, %	Control	Treatment	Control	Treatment	Control	Treatment
Corn, yellow dent	35.71	35.71	40.57	39.57	45.69	44.69
Soybean meal, 47.5% CP	13.5	13.5	18.00	18.00	23.00	23.00
Soy protein concentrate	3.12	3.12	2.50	2.50	3.00	3.00
Soybean oil	3.62	3.62	4.00	4.00	1.66	1.66
Corn DDGS, 7%	0.00	0.00	0.00	0.00	10.00	10.00
Plasma, spray-dried	6.5	6.5	2.50	2.50	0.00	0.00
Selenium premix	0.05	0.05	0.05	0.05	0.05	0.05
Blood meal, spray-dried	1.00	1.00	1.00	1.00	0.00	0.00
Whey, dried	25.00	25.00	25.00	25.00	10.00	10.00
Fish meal, menhaden	4.00	4.00	4.00	4.00	4.00	4.00
Lactose	5.00	5.00	0.00	0.00	0.00	0.00
Limestone	1.17	1.17	0.94	0.94	0.93	0.93
Monocalcium Phosphate 21%	0.11	0.11	0.21	0.21	0.21	0.21
Trace Mineral premix[1]	0.13	0.13	0.15	0.15	0.15	0.15
Vitamin premix[2]	0.25	0.25	0.25	0.25	0.25	0.25
Salt	0.25	0.25	0.25	0.25	0.47	0.47
Natuphos 600[3]	0.10	0.10	0.10	0.10	0.10	0.10
L-lysine-HCl	0.28	0.28	0.26	0.26	0.28	0.28
L-threonine	0.04	0.04	0.04	0.04	0.03	0.03
DL-methionine	0.18	0.18	0.15	0.15	0.15	0.15
L-tryptophan	0.00	0.00	0.03	0.03	0.03	0.03
Fiber premix[4]	0.00	0.00	0.00	1.00	0.00	1.00
ME, Kcal/lb	1594	1594	1592	1577	1543	1527
Crude Protein, %	23.26	23.26	22.27	22.19	23.57	23.49
Lysine, %	1.79	1.79	1.62	1.62	1.54	1.53
SID Lysine, %	1.64	1.64	1.48	1.48	1.37	1.37
SID Threonine, %	0.96	0.96	0.87	0.87	0.81	0.81
SID Tryptophan, %	0.28	0.28	0.28	0.28	0.26	0.26
SID Methionine, %	0.50	0.50	0.47	0.47	0.47	0.46
SID Methionine + Cysteine, %	0.90	0.90	0.81	0.81	0.76	0.76
SID Valine, %	1.07	1.07	0.98	0.97	0.97	0.96
SID Isoleucine, %	0.84	0.84	0.82	0.82	0.87	0.86
SID Leucine, %	1.88	1.88	1.76	1.75	1.58	1.57
Calcium, %	0.94	0.94	0.88	0.88	0.78	0.78
Phosphorous, %	0.64	0.64	0.64	0.64	0.63	0.62
Available Phosphorous, %	0.43	0.43	0.42	0.42	0.35	0.35

[1]Trace mineral premix provided the following guaranteed minimums per kg diet: iron, 84.7 mg; zinc, 84.7 mg; manganese, 10.5 mg; copper, 7.87 mg; iodine, 0.32 mg.

[2]Vitmain premix provided the following guaranteed minimums per kg diet: vitamin A, 9000 IU; vitamin E, 187 IU; vitamin K (hetrazeen), 2.62 mg; vitamin B_1, 1.857 mg; vitamin B_{12}, 17.25 µg; riboflavin, 5.25 mg; d-pantothenic acid, 11.25 mg; niacin, 18.75 mg.

[3]Contains 600 U/g of phytase activity

On day 4 post-weaning, all pigs were moved to individual housing, and were given the Phase 2 diet. Phase 3 nursery diets began on d 11 and were fed until the conclusion of the study (d 21). For both Phase 2 and 3 diets, the basal diet was made without supplemental dextrin, and pigs getting dextrin received the basal diet blended with supplemental dextrin at a level of 1% of the diet, whereas pigs not receiving supplemental dextrin had an extra 1% corn blended into the diet (Table 3.1).

3.2.2 Animal growth and performance

Feed intake and animal weights were measured at different life stages. BW, average daily gain (ADG), average daily feed intake (ADFI), and gain:feed (G:F) were determined on a pen basis from d 0 to d 4 and on an individual animal basis from d 4 to d 21.

3.2.3 Sample collection

On d 0 and d 21, 8 and 31 pigs were euthanized, respectively, using carbon dioxide stunning followed by exsanguination. Intestinal tissue was collected from the ileum and cecum. A 2-inch cross section was taken from the proximal ileum (10 inches from the ileal-cecal junction), rinsed with phosphate-buffered saline (PBS), and then placed in 30 mL of 10% neutral buffered formalin for subsequent histological analyses. Proximal ileal and cecal tissue were scraped and flash frozen in 1 mL of Invitrogen™ TRIzol® reagent (ThermoFisher Scientific; Waltham, MA, USA) for subsequent isolation of mRNA. Digesta and mucosal swab samples (from microbiota analysis) were taken from the distal ileum, cecum, and proximal large intestine (PLI), placed on ice and transported to the laboratory. Mucosal swab samples were taken by cleaning the tissue with PBS then using a cotton swab to wipe a 2 inch by 2 inch area. Fecal samples were also taken, placed in a sterile bag, transported on ice and stored at -20°C until short-chain fatty acid (SCFA) analysis.

Table 3.2 Analyzed composition of post-weaning swine diets

	Phase 1	Phase 2		Phase 3	
Nutrient[1]		No Fiber	Fiber	No Fiber	Fiber
Nitrogen, %	3.81	3.43	3.67	3.54	3.49
Metabolizable Energy, Kcal/kg	4152	3506	4160	4037	4019
Dry Matter, %	88.42	86.97	87.75	86.8	87.25
Ash, %	5.2	5.37	5.27	4.67	4.6
Phosphorous, %	0.667	0.703	0.685	0.739	0.628
NDF, %	1.83	5.15	3.95	10.26	13.29
ADF, %	-0.765	1.86	0.114	3.12	5.73

[1] Analyzed at Purdue University

3.2.4 Histology

Ileal cross sections that were placed in formalin were prepared and imaged at the Purdue University Histology Lab in the College of Veterinary medicine. Digital pictures of the slides were taken with MotiConnect software (version 1.5.9.1 ; Motic China Group Co. Ltd.; Xiamen, China). The digitized slides were then analyzed using ImageJ 1.51k measurement software (LOCI, University of Wisconsin). Six villi and six crypts were measured for each pig. The villi were measured from the tip to the base of the villus. Crypts were measured from the base of the villus to the bottom of the crypt region.

3.2.5 Short-chain Fatty Acid (SCFA) Analysis

Fecal samples were removed from the freezer and thawed in a refrigerator. Thawed feces were diluted with a recorded amount of deionized (DI) water until they reached the desired liquid consistency, and then mixed with 2 mL of 25% metaphosphoric acid and placed into tubes. Each sample was divided into 2 tubes so that each sample could be run in duplicate. Samples were then vortexed for 10 seconds, followed by centrifugation (15,120 x *g* for 10 minutes) after which the supernatant was collected. The samples were centrifuged again at 15,120 x *g* for 10 minutes and the supernatant was collected and poured into a 12-cc syringe, which connected to a paper filter (0.45 μm pore size). The supernatant was passed through the filter and the filtrate was collected. A stock solution with 87.4 μM/L acetic acid, 26.76 μM/L propionic acid, 7.607 μM/L butyric acid, 8.789 μM/L isobutyric acid, 1.864 μM/L valeric acid, and 1.835 μM/L isovaleric acid was mixed as a standard to normalize the samples. The Gas Chromatograph (3900 CP-8400, Varian Medical

Systems™, Palo Alto, CA, USA) installed with a fused silica capillary column (0.25mm x 0.25µm x 30 m), was loaded with a blank containing a 1:5 dilution of 25% metaphosphoric acid and a standard containing a 1:5 dilution of the standard stock, followed by the experimental samples. The sample analysis was run using Galaxie analysis software (Galaxie, Varian Medical Systems™, Palo Alto, CA, USA).

3.2.6 Microbiome sample preparation and sequencing

Total community DNA, extracted using the Qiagen MagAttract PowerMicrobiome DNA/RNA Kit, was used for the construction of a 16S rRNA gene library following a standardized protocol (Kozich et al., 2013). Briefly, Illumina indexed amplicons were created using PCR amplification of the V4 region of bacterial 16S rRNA gene. Amplification success was determined through gel electrophoresis as a quality check. No bands were observed in the negative control samples in which water was used as the DNA template. Amplified DNA was normalized using a SequalPrep Normalization Plate (Invitrogen) and pooled into a single library for each 96-well plate. Library concentration was determined using the KAPA Library Quantification Kit (Roche) and library average fragment length was determined using the Bioanalyzer (Agilent) with a high sensitivity kit. Following the confirmation of proper DNA concentration, the pooled samples, mock community, and water were sequenced (Illumina, MiSeq v2 kit, 500 cycle). Sequences were demultiplexed according to oligonucleotide bar code sequence with Illumina software. Sequences were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) (BioProject PRJNA687488; BioSamples SAMN17145353 – SAMN17145574).

3.2.7 Bioinformatics for microbiome sequencing

Raw reads were analyzed using the mothur (Schloss et al., 2009; v 1.39.3) MiSeq analysis standard operating procedure. The general pipeline for mothur is as follows: make contiguous sequences (contigs) from raw reads, align contigs to reference sequences (Quast et al., 2012; SILVA database release 132), screen and filter sequences to remove low quality reads (ambiguous bases allowed = 0, maximum read length = 275, homopolymers allowed = 8), group sequences based on sequence similarity, classify sequences with reference to known taxonomic classifications (Cole et al., 2013; RDP training set 16), cluster sequences, and run diversity metrics.

3.2.8 Statistical Analysis

For pig growth performance during the pre-weaning period, intestinal morphology, and intestinal gene expression (n=8), all data were analyzed with two treatment groups, with and without supplemental dextrin. For pigs euthanized on day 21, the pig was used as the experimental unit (n=31) for growth performance, gene expression, histology and SCFA concentrations. Data were analyzed using PROC GLM in SAS 9.4 (SAS Inst. Inc., Cary, NC). All growth performance data were analyzed as a 2x2 factorial arrangement with or without supplemental fiber pre-weaning and with or without supplemental fiber post-weaning. Values were considered significant at $P \leq 0.05$ and a trend at $0.05 < P \leq 0.10$. Any data point greater than 2.5 standard deviations from the average was removed and not included in statistical analysis. Four additional pigs, one per treatment, were removed from the final data set for analysis due to poor health. Because the animals cannot be blocked by microbial community, for microbiome data the treatments were not analyzed using a 2x2 factorial design and instead as whole treatment groups per animal (pre-weaning treatment / post-weaning treatment: Fiber/Fiber, No Fiber/Fiber, Fiber/No Fiber, and No Fiber/No Fiber). Alpha and beta diversity were calculated using mothur (Schloss et al., 2009; v 1.39.3) and analyzed with R statistical software (R Core Team, 2013; v 1.1.423). Alpha diversity was visualized with box-and-whisker plots using the R package "ggplot2" and differences were determined using Mann-Whitney-Wilcoxon or Kruskal-Wallis with post-hoc pairwise comparisons where appropriate. Beta diversity was calculated using Bray-Curtis dissimilarity and Jaccard from the R package "vegan" and visualized with non-metric dimensional scaling (NMDS) plots using ggplot2. Dispersion was calculated, using "betadisper" from the "vegan" package, for Bray Curtis and Jaccard indices to determine the homogeneity of the treatment groups while perMANOVA was used to determine differences in variance of the treatment groups. Metastats was used to determine statistical significance between treatment groups in mothur (Schloss et al., 2009; v 1.39.3). Multiple test corrections, calculated using Benjamini-Hochberg, were completed in R (R Core Team, 2013; v 1.1.423) for all necessary statistical tests and corrections are labeled as "q-value". All scripts used to produce mothur and R output, as well as the accompanying data, can be found at https://github.com/clwickwa/16S_Analysis/SwineFiber.

3.3 Results

3.3.1 Animal Growth and Performance

Dextrin supplementation had limited to no impact on animal feed intake, weight gain, or feed efficiency during the study period. There was no difference in ADG while pigs were nursing ($P > 0.05$; Table 3.3), but on day 0, pigs receiving supplemental dextrin tended to weigh less than those not receiving dextrin ($P = 0.087$). This tendency became significant four days after weaning, and pigs that received dextrin following weaning had decreased body weight ($P = 0.022$; Table 3.3). From d 4 to d 11, there were no differences in ADG or ADFI ($P > 0.05$; Table 3.3), however there was a tendency for pigs receiving dextrin after weaning to have decreased G:F d 4 to d 11 ($P = 0.056$; Table 3.3). From d 11 to 21 there was a tendency in ADG for an interaction between pre- and post-weaning dextrin supplementation ($P = 0.085$; Table 3.3), but no differences in ADFI, G:F, or d 21 BW ($P > 0.05$; Table 3.3) were observed. There were no differences between treatments for ADG, ADFI, or G:F from d 0-21 ($P > 0.05$; Table 3.3). Overall, dextrin supplementation tended to reduce BW for the first 4 day period post-weaning, but there were no differences in any animal growth performance measures during the pre-weaning period, after the first 4 days post-weaning, or the overall post-weaning period due to dextrin supplementation.

Table 3.3 Effect of dextrin supplementation before and after weaning on growth and development

						Probability, P<		
Animal Parameter	NF/NF	F/NF	NF/F	F/F	SE	Pre-wean Main Effect	Post-wean Main Effect	Pre-weaning x Post-weaning
Initial Wt, kg	3.26	3.00	3.13	3.25	0.138	0.587	0.648	0.158
d -3 Wt, kg	5.68	5.53	5.34	5.58	0.149	0.771	0.317	0.179
Day -14 - 0								
ADG, kg/d	0.223	0.224	0.199	0.215	0.014	0.507	0.216	0.533
d 0 Wt, kg	6.38	6.13	5.91	6.26	0.179	0.778	0.332	0.087
Day 0 - 4								
ADG, kg/d	0.105	0.098	0.101	0.044	0.040	0.403	0.445	0.514
ADFI, kg/d	0.139	0.132	0.165	0.119				
d 4 Wt, kg	6.80	6.52	6.06	6.44	0.188	0.769	0.022	0.058
Day 4 - 11								
ADG, kg/d	0.189	0.142	0.118	0.077	0.049	0.356	0.164	0.951
ADFI, kg/d	0.264	0.367	0.243	0.202	0.070	0.651	0.183	0.297
G:F	0.532	0.444	0.059	0.215	0.179	0.849	0.056	0.486
d 11 Wt, kg	7.91	7.51	7.15	6.98	0.456	0.508	0.146	0.792
Day 11- 21								
ADG, kg/d	0.220	0.323	0.303	0.216	0.055	0.876	0.831	0.085
ADFI, kg/d	0.425	0.470	0.540	0.382	0.067	0.377	0.824	0.120
G:F	0.487	0.645	0.533	0.476	0.089	0.562	0.481	0.219
d 21 Wt, kg	10.33	10.58	10.51	9.36	0.935	0.615	0.559	0.432
Overall Post-Weaning Day 4 - 21								
ADG, kg/d	0.196	0.226	0.233	0.162	0.048	0.587	0.697	0.276
ADFI, kg/d	0.364	0.401	0.426	0.312	0.053	0.451	0.795	0.148
G:F	0.475	0.408	0.509	0.370	0.075	0.974	0.186	0.623

[1]Sample size: preweaning NF/NF (n = 10), F/NF (n = 10), NF/F (n = 9), F/F (n = 10); postweaning: NF/NF (n = 8), F/NF (n = 8), NF/F (n = 7), F/F (n = 8)

3.3.2 Histology

Pigs euthanized prior to weaning on d 0 had no differences in villus height, crypt depth, or villus height to crypt depth ratio ($P > 0.1$), however, differences were observed in intestinal morphology by the end of the study on d 21. Providing supplementary dextrin to pigs prior to weaning tended to decrease crypt depth ($P = 0.0904$), while villus height was unchanged ($P > 0.1$ Table 3.4), leading to a tendency ($P = 0.093$) for an increased villus height to crypt depth ratio in animals given fiber during the pre-weaning period. The addition of fiber after weaning resulted in no difference in villus height, crypt depth, or villus height:crypt depth ratio ($P > 0.05$; Table 3.4).

3.3.3 Short Chain Fatty Acid (SCFA) Analysis

In general, addition of dextrin to the diet caused an increase in the concentration of short chain fatty acids in the feces, especially for pigs that received dextrin during the pre-weaning period (Table 3.5). For pigs euthanized on d 21 post weaning, dextrin supplementation resulted in a pre-weaning treatment x post-weaning treatment interaction for the total amount SCFAs (P = 0.054). Pigs that received supplemental dextrin at any point had greater concentrations of total SCFAs compared to pigs that never received supplemental dextrin. Acetate and propionate concentrations (mmol/L) were not different among treatment groups ($P > 0.10$). There was an interaction ($P = 0.007$) of pre- and post-weaning treatment for butyrate concentrations which was explained by pigs receiving supplemental dextrin at any point having increased butyrate concentrations. However, pigs that received supplemental dextrin only prior to weaning or after weaning had greater concentrations of butyrate when compared to pigs that received dextrin for the entire study (Table 3.5). A pre-weaning treatment x post-weaning treatment interaction was observed for valerate concentrations. Pigs fed supplemental dextrin pre-weaning but not post-weaning had greater concentrations of valerate compared to all other treatment groups ($P = 0.045$) and pigs fed supplemental dextrin for the entirety of the study had decreased fecal valerate concentrations. Pigs receiving supplemental dextrin prior to weaning had decreased isobutyrate (P = 0.050) and a tendency for decreased isovalerate ($P = 0.058$) concentrations compared to other treatment groups (Table 3.5).

When SCFA data were analyzed as a percentage of total SCFAs in the feces, acetate accounted for a larger proportion of SCFAs for pigs receiving supplemental dextrin post-weaning ($P = 0.047$; Table 3.5). Propionate as a percentage of total SCFAs did not differ among treatment groups ($P \geq 0.262$; Table 3.5). There was an interaction between pre- and post-weaning dextrin supplementation for butyrate concentrations, with pigs being fed supplemental dextrin only prior to weaning having a greater proportion of butyrate (as a percentage of total SCFAs) than all other treatment groups ($P = 0.029$; Table 3.5). Pigs receiving supplemental dextrin post-weaning had a decreased valerate concentration as a percentage of total SCFAs ($P = 0.038$; Table 3.5). Decreases in isobutyrate ($P = 0.040$) and a tendency for isovalerate ($P = 0.051$) as a percentage of total SCFAs were observed in pigs that received supplemental dextrin pre-weaning compared to pigs that did not receive supplemental dextrin pre-weaning (Table 3.5).

3.3.4 Microbiota diversity measures

A total of 338 operational taxonomic units (OTUs) were observed in this study. The predicted number of OTUs (Chao Index) was 100-300 OTUs from mucosal swabs and luminal contents in a majority of the samples from the ileum. In comparison, the Chao indices for mucosal swabs and luminal contents from cecal and PLI samples was higher, between 200-500 OTUs. No differences ($P > 0.10$) were observed in microbial community alpha diversity (Chao or Shannon indices) based on diet or sample type (luminal or mucosal).

As expected, ileal microbial communities were clearly distinct from cecal and PLI communities for both Bray-Curtis and Jaccard indices (PERMANOVA; $P < 0.05$, Table 3.6). Additionally, mucosal communities were significantly different from digesta communities in post-weaning pigs (PERMANOVA; $P < 0.05$, Table 3.6). Luminal communities had an increased relative abundance of *Coprococcus*, *Butyricimonas,* and *Anaerovibrio*, while mucosal communities had an increased relative abundance of *Mucispirillum, Desulfovibrionaceae,* and *Novosphingobium Helicobacter*, and *Prevotella* ($q < 0.05$; Table 3.8).

Supplementation of dextrin in the diet altered the community composition on d 21 post-weaning, as a main effect, in both mucosal and luminal samples (PERMANOVA; $q < 0.05$, Table 3.6C-D). Pairwise PERMANOVA testing revealed that nearly all diet groups were distinct from each other ($q < 0.05$, Table 3.7) in the mucosal samples, but diet explained less than 10% of the variation, indicating that the differences were not dramatic. Pairwise testing revealed there was a tendency for dextrin treatments to result in a shift in the luminal intestinal microbial communities (PERMANOVA; $q < 0.1$, Table 3.7). No difference in intestinal microbial communities of mucosal swabs or luminal contents were observed due to supplemental dextrin during the pre-weaning period only, likely due to the low number of samples collected (n = 4 per treatment) and the low number of samples with sufficient number of high quality sequences (n = 2 in ileum groups).

3.3.5 Microbiome composition

Immediately following weaning there was a large degree of inter-individual variance, but the most dominant genera in the piglets overall were *Lactobacillus* and *Prevotella* (Fig 1). The most abundant genera also included bacteria that may induce inflammation and disease, such as *Fusobacterium*, especially in the mucosal samples (Figure 3.1B). Shifts in the relative abundance

of genera after the pre-weaning period were not significant; however, there were numerical shifts in *Lactobacillus*, *Clostridium sensu stricto*, and *Prevotella* between pigs due to supplemental dextrin that likely deserve future consideration. In the ileal digesta and mucosa, pigs that received supplemental dextrin had a numerically decreased abundance of *Lactobacillus*.

Table 3.4 Interactive effects of dextrin supplementation pre- and post-weaning on villus height, crypt depth, and villus height:crypt depth on d21 post-weaning

Diet	NF/NF	F/NF	NF/F	F/F	SE	Probability, P< Pre-wean	Post-wean	Pre-weaning x Post-weaning
Villus Height, μm	378	367	364	336	20.49	0.336	0.275	0.675
Crypt depth, μm	346	314	325	295	18.01	0.09	0.266	0.975
Villus:Crypt	1.074	1.222	1.111	1.257	0.089	0.093	0.671	0.994

Table 3.5 Main and interaction effects of dextrin supplementation pre- and post-weaning on SCFA concentrations and percentages

Diet[1]	NF/NF	NF/F	F/NF	F/F	SE	Probability, P< Pre-wean	Post-wean	Pre-weaning
Total SCFA, mmol/L	144	182	180	164	14.0	0.507	0.401	0.054
Acetate, mmol/L	73	96	87	86	7.68	0.773	0.141	0.115
Acetate, % of total	51	52	48	53	1.54	0.595	0.047	0.216
Propionate, mmol/L	38	47	44	43	4.84	0.830	0.452	0.266
Propionate, % of total	26	25	25	26	1.06	0.486	0.966	0.262
Butyrate, mmol/L	24	30	37	28	2.70	0.057	0.497	0.007
Butyrate, % of total	17	17	21	16	1.02	0.070	0.047	0.029
Valerate, mmol/L	6.1	6.3	9.7	5.9	0.990	0.096	0.072	0.045
Valerate, % of total	4.2	3.7	5.4	3.5	0.555	0.314	0.038	0.190
Isobutyrate, mmol/L	1.2	1.4	0.88	0.83	0.237	0.050	0.746	0.571
Isobutyrate, % of total	0.87	0.92	0.50	0.50	0.190	0.040	0.905	0.878
Isovalerate, mmol/L	1.2	1.4	0.77	0.85	0.304	0.058	0.745	0.961
Isovalerate, % of total	0.95	0.99	0.44	0.50	0.253	0.051	0.832	0.949

[1]Sample size: NF/NF (n = 8), NF/F (n = 7), F/NF (n = 8), F/F (n = 8)

Three weeks after weaning, the major members of the community included *Lactobacillus*, *Prevotella*, *Campylobacter*, unclassified *Veillonellaceae, Megasphaera*, and *Escherichia*. There were statistically significant shifts in OTUs (metastats, q <0.05) of the cecal and PLI microbial communities 21 d post weaning, mainly in less abundant OTUs (abundance of less than 1% of the community) (Table 3.8).

When comparing groups with different dextrin supplementation during the preweaning period (NF/NF vs F/NF; Table 3.8A), many OTUs were enriched when no supplemental dextrin was added in the post-weaning period (NF/NF, F/NF. Conversely, when dextrin was fed during the post-weaning period but differentially fed during the pre-weaning period (NF/F vs F/F), there were no differentially abundant OTUs. Some OTUs enriched in NF/NF compared to F/NF included *Coprococcus*, *Butyricimonas*, *Mogibacterium*, *Anaerovibrio* and *Paraprevotella* (Table 3.8A). One OTU, identified as unclassified *Desulfovibrionaceae,* was increased in the pigs fed no supplemental dextrin compared to pigs fed supplemental dextrin during pre-weaning - NF/NF vs F/NF (Table 3.8A) and NF/NF vs F/F (Table 3.8B).

When comparing groups that differed in supplemental dextrin in the post-weaning period (F/NF vs F/F (Table 3.8C) or NF/NF vs NF/F (Table 3.8D)), the result was inconsistent. There was a single differentially abundant OTU when dextrin was not supplemented during the pre-weaning period but was altered during the post-weaning period compared to the groups differentially fed supplemental fiber pre-weaning (NF/NF vs NF/F (Table 3.8D) and F/NF vs NF/F (Table 3.8E)).

Table 3.6 PERMANOVA main effects of dextrin supplementation and sampling region on bacterial community beta diversity

A. Pre-weaning mucosal

	Df	SumsOfSqs	MeanSqs	F.Model	R^2	Pr(>F)
Diet	1	0.3162	0.3162	1.08164	0.05289	0.3579
Region	2	0.8974	0.4487	1.53488	0.1501	0.057
Diet:Region	2	0.3801	0.19007	0.65019	0.06358	0.9099
Residuals	15	4.3851	0.29234	0.73343		
Total	20	5.9788	1			

B. Pre-weaning contents

	Df	SumsOfSqs	MeanSqs	F.Model	R^2	Pr(>F)
Diet	1	0.198	0.19796	0.64491	0.03001	0.8002
Region	2	1.3707	0.68535	2.23274	0.20782	0.0058*
Diet:Region	2	0.4228	0.21139	0.68867	0.0641	0.8585
Residuals	15	4.6043	0.30696	0.69807		
Total	20	6	1			

C. Post-weaning mucosal

	Df	SumsOfSqs	MeanSqs	F.Model	R^2	Pr(>F)
Diet	3	1.5014	0.50046	2.5596	0.07915	0.0001*
Region	2	2.7358	1.36791	6.9964	0.14422	0.0001*
Diet:Region	6	1.0459	0.17432	0.8916	0.05514	0.7551
Residuals	70	13.686	0.19552	0.72149		
Total	81	18.9693	1			

D. Post-weaning contents

	Df	SumsOfSqs	MeanSqs	F.Model	R^2	Pr(>F)
Diet	3	1.2151	0.40504	2.1184	0.0628	0.0005*
Region	2	3.3711	1.68555	8.8156	0.17422	0.0001*
Diet:Region	6	0.9974	0.16623	0.8694	0.05154	0.7711
Residuals	72	13.7664	0.1912	0.71144		
Total	83	19.35	1			

Table 3.7 Pairwise PERMANOVA results for dextrin supplementation and sampling region

A. Pre-weaning contents, region				
	F.Model	R^2	p.value	q
cecum vs ileum	3.7046035	0.25193495	0.0028	0.0042*
cecum vs colon	0.1975277	0.01391282	0.9807	0.9807
ileum vs colon	3.8896915	0.26123385	0.0014	0.0042*
B. Post-weaning mucosal, diet				
	F.Model	R^2	p.value	q
NF/NF vs F/NF	2.020038	0.04807321	0.0256	0.03840*
NF/NF vs F/F	1.399386	0.03224438	0.1412	0.1412
NF/NF vs NF/F	3.359982	0.08123751	0.0002	0.00120*
F/NF vs F/F	1.824334	0.04361897	0.0416	0.04992*
F/NF vs NF/F	2.494069	0.06479099	0.0086	0.01720*
F/F vs NF/F	2.601212	0.06406735	0.0016	0.00480*
C. Post-weaning mucosal, region				
	F.Model	R^2	p.value	q
cecum vs ileum	9.486399	0.15947173	0.0001	0.00015*
cecum vs colon	2.012331	0.03353196	0.0132	0.0132
ileum vs colon	8.961158	0.15198409	0.0001	0.00015*
D. Post-weaning contents, diet				
	F.Model	R^2	p.value	q
NF/NF vs F/NF	1.87048	0.04263642	0.0324	0.0648
NF/NF vs F/F	1.691441	0.04057046	0.0683	0.10245
NF/NF vs NF/F	2.207446	0.05490143	0.0237	0.0648
F/NF vs F/F	1.499668	0.03447538	0.1138	0.1138
F/NF vs NF/F	1.921116	0.04582693	0.0303	0.0648
F/F vs NF/F	1.622617	0.0409518	0.09	0.108
E. Post-weaning contents, region				
	F.Model	R^2	p.value	q
cecum vs ileum	11.05453	0.17531698	1.00E-04	0.00015*
cecum vs colon	2.960038	0.04855703	5.00E-04	0.00050*
ileum vs colon	11.77578	0.18464345	1.00E-04	0.00015*

[1]Main effect significant results only

Table 3.8 Differentially abundant operational taxonomic units (OTUs) between diet groups post-weaning (q < 0.05). No distinguishing OTUs found between the any of the ileal samples for either mucosal swabs or digesta, nor pre-weaning pigs.

A. No Fiber/No Fiber vs. Fiber/No Fiber								
OTU	Genus	NF/NF	NF/NF SE	F/NF	F/NF SE	q value	Region	Sample type
Otu051	*Coprococcus*	0	0	0.000366	0.000335	0.0181	cecum	digesta
Otu063	*Butyricimonas*	0	0	0.000761	0.000761	0.0181	cecum	digesta
Otu066	*Bilophila*	0	0	0.000282	0.000282	0.0181	cecum	digesta
Otu069	*Mogibacterium*	0	0	0.002113	0.002113	0.0181	cecum	digesta
Otu092	*Anaerovibrio*	0	0	0.000282	0.000282	0.0181	cecum	digesta
Otu110	*Paraprevotella*	0	0	0.002366	0.002366	0.0181	cecum	digesta
Otu153	*Psychrobacter*	0	0	0.000592	0.000592	0.0181	cecum	digesta
Otu022	*Streptococcus*	0.000479	0.000231	0.00924	0.003507	0.0422	cecum	digesta
Otu061	*Pseudoscardovia*	0.007409	0.004607	0.000366	0.000127	0.0422	cecum	digesta
Otu051	*Coprococcus*	0	0	0.000535	0.000535	0.0339	colon	digesta
Otu092	*Anaerovibrio*	0	0	0.000423	0.000423	0.0339	colon	digesta
Otu110	*Paraprevotella*	0	0	0.007606	0.007606	0.0339	colon	digesta
Otu152	*Eggerthella*	0	0	0.000366	0.000366	0.0339	colon	digesta
Otu066	*Bilophila*	0	0	0.000254	0.000254	0.0314	colon	mucosal
Otu095	*Desulfovibrionaceae* unc.[1]	0.000676	0.000586	0	0	0.0314	colon	mucosal
Otu118	*Lactococcus*	0	0	0.000761	0.000761	0.0314	colon	mucosal
Otu122	*Victivallis*	0	0	0.000282	0.000282	0.0314	colon	mucosal
B. No Fiber/No Fiber vs. Fiber/Fiber								
OTU	Genus	NF/NF	NF/NF SE	F/F	F/F SE	q value	Region	Sample type
Otu028	*Clostridium sensu stricto*	0	0	0.001099	0.000635	0.0412	cecum	digesta
Otu051	*Coprococcus*	0	0	0.001859	0.001608	0.0412	cecum	digesta
Otu092	*Anaerovibrio*	0	0	0.000282	0.000282	0.0412	cecum	digesta
Otu058	*Mucispirillum*	0.000592	0.000259	0.002648	0.000692	0.0292	cecum	mucosal
Otu092	*Anaerovibrio*	0	0	0.000225	0.000225	0.0292	cecum	mucosal
Otu095	*Desulfovibrionaceae* unc.[1]	0.000254	0.000197	0	0	0.0292	cecum	mucosal
Otu204	*Novosphingobium*	0	0	0.000451	0.000451	0.0292	cecum	mucosal
C. Fiber/No Fiber vs. Fiber/Fiber								
OTU	Genus	F/NF	F/NF SE	F/F	F/F SE	q value	Region	Sample type
Otu075	*Corynebacterium*	0.000282	0.000189	0	0	0.0342	cecum	digesta
Otu110	*Paraprevotella*	0.002366	0.002366	0	0	0.0342	cecum	digesta
Otu153	*Psychrobacter*	0.000592	0.000592	0	0	0.0342	cecum	digesta
Otu173	*Rhodobacter*	0.001099	0.001099	0	0	0.0342	cecum	digesta
Otu058	*Mucispirillum*	0.000648	0.00032	0.002962	0.000404	0.0259	colon	digesta
Otu059	*Subdoligranulum*	0	0	0.001465	0.001433	0.0223	colon	mucosal

Table 2.8 Continued

C. Fiber/No Fiber vs. Fiber/Fiber (continued)

OTU	Genus	F/NF	F/NF SE	F/F	F/F SE	q value	Region	Sample type
Otu091	*Actinobacteria* unc.[1]	0	0	0.000507	0.000272	0.0223	colon	mucosal
Otu118	*Lactococcus*	0.000761	0.000761	0	0	0.0223	colon	mucosal
Otu134	*Reyranella* unc.	0.00093	0.000505	0	0	0.0223	colon	mucosal
Otu204	*Novosphingobium*	0	0	0.000338	0.000338	0.0223	colon	mucosal
Otu230	*Ezakiella*	0	0	0.000338	0.000338	0.0223	colon	mucosal
Otu054	*Enterobacteriaceae* unc.[1]	0.010987	0.009001	0.000113	7.40E-05	0.0382	colon	mucosal

D. No Fiber/No Fiber vs. No Fiber/Fiber

OTU	Genus	NF/NF	NF/NF SE	NF/F	NF/F SE	q value	Region	Sample type
Otu058	*Mucispirillum*	0.000592	0.000259	0.006665	0.001313	0.0234	cecum	mucosal

E. Fiber/No Fiber vs. No Fiber/Fiber

OTU	Genus	F/NF	F/NF SE	NF/F	NF/F SE	q value	Region	Sample type
Otu058	*Mucispirillum*	0.00058	0.00024	0.006665	0.001313	0.0410	cecum	mucosal

[1] Suffix “unc.” refers to OTUs that were not classified to the genus level – taxonomic identifier is at the family level

Many OTUs with increased or decreased abundance were found when dextrin was fed during the pre-weaning period but differed in the post-weaning period (F/NF vs F/F (Table 3.8C)). While there were many differences in OTU abundance in both F/NF and F/F treatment groups, potentially beneficial bacteria were enriched in both treatment groups. When comparing the groups that differed in dextrin feeding during both periods (NF/NF vs F/F, Table 3.8B), the differential abundance pattern was similar to F/NF vs NF/NF, with *Coprococcus* and *Anaerovibrio* enriched in the dextrin fed animals, while the no fiber group was enriched with unclassified *Desulfovibrionaceae*.

3.4 Discussion

The stress of weaning and change in diet are known to cause significant shifts in the intestinal microbial community. We sought to determine if diet supplementation with the soluble fiber dextrin could cause this shift in microbial community to occur earlier in the life of the piglet, in preparation for weaning and diet change. Often the stress at weaning requires the use of prophylactic or therapeutic antibiotics and reducing the need for antimicrobial use in animals is included in the One-Health approach to reduce antibiotic resistance in pathogens. The swine industry, as with all other users of antibiotics, is working to develop different means of lowering the need for antimicrobials through supplementation of various pathogen-suppressing feed additives.

Several researchers have reported the many benefits that are accompanied with adding fiber to swine diets (Cleave, 1968; Trowell, 1972). Dietary fiber has been reported by (Jenkins et al., 2002) to increase the concentration of SCFAs which are capable of being used by the enterocytes as an energy source.

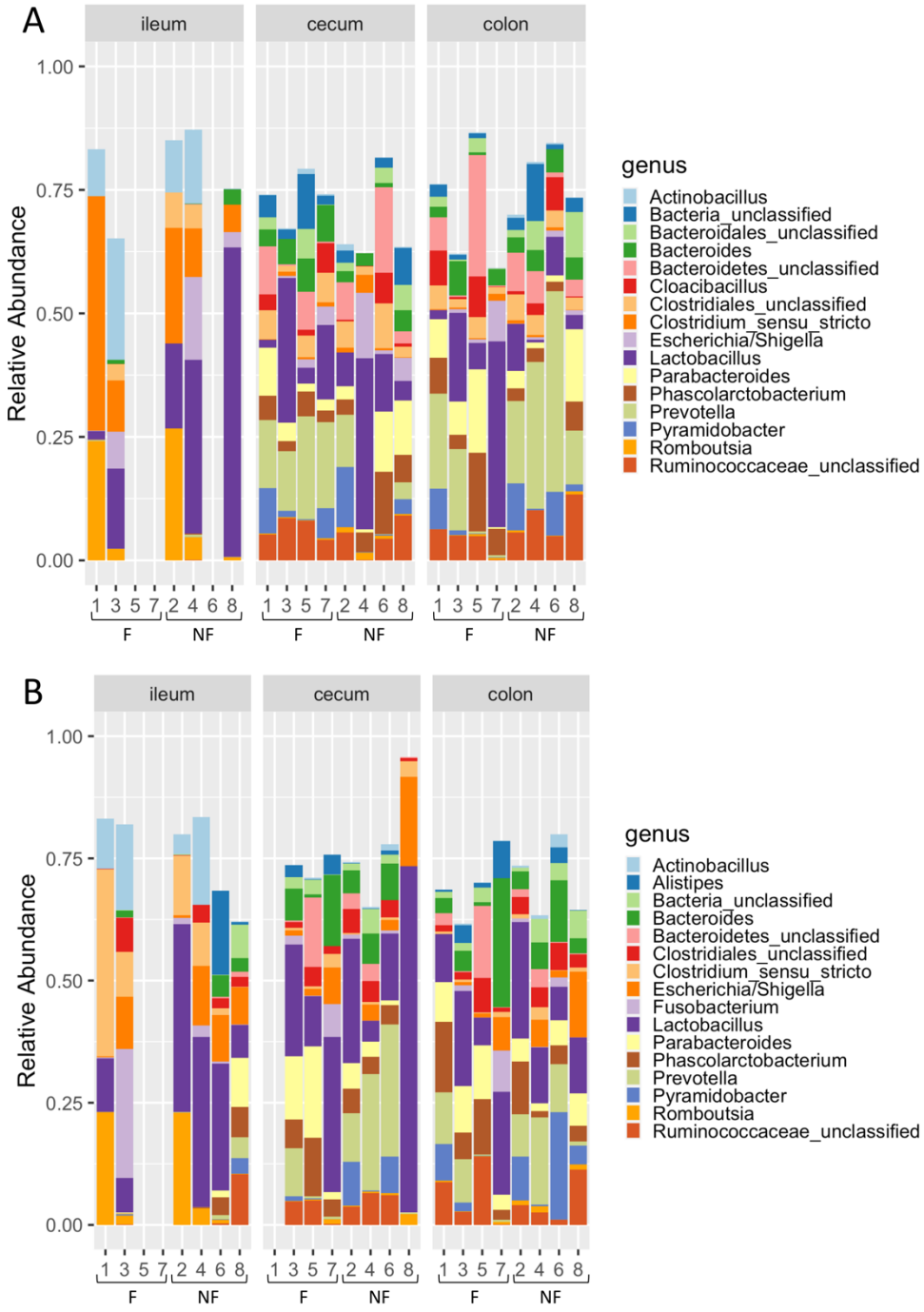

Figure 3.1 Top 16 most abundant genera of pre-weaning swine gut bacterial communities

Relative abundance of genera (y-axis) for individual pigs at the end of the suckling period. Both the luminal (A) and mucosal (B) microbial communities were characterized. Fiber (F) n = 4, No Fiber (NF) n = 4. Samples that had low sequence counts were not included in the analysis. PLI = proximal large intestine

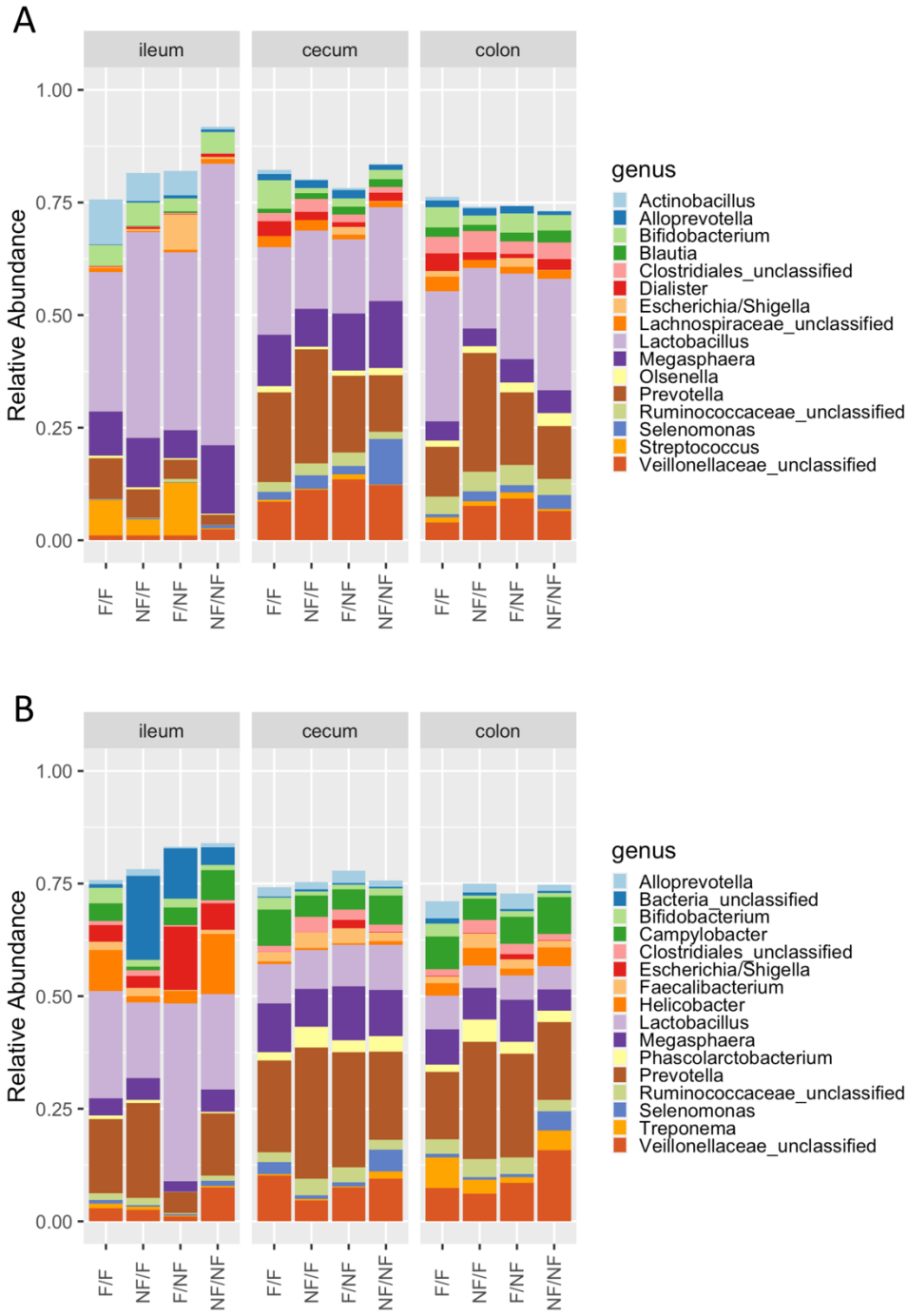

Figure 3.2 Top 16 most abundant genera of post-weaning swine gut bacterial communities
Relative abundance of genera (y-axis) for diet group averages at the end of the experimental period. Both the luminal (A) and mucosal (B) microbial communities were characterized. Pre- and post-weaning diet treatments are indicated: fiber pre-weaning and fiber post-weaning (F/F) n = 8, no fiber pre-weaning and fiber post-weaning (NF/F) n = 7, fiber pre-weaning and no fiber post-weaning (F/NF) n = 8, no fiber pre-weaning and no fiber post-weaning (NF/NF) n = 8. Samples that had low sequence counts were not included in the analysis.
PLI = proximal large intestine

SCFAs have been reported to boost the efficiency of the immune system by increasing the amount of natural killer cells (Pratt et al., 1996) and liberates glutamine to be used as an energy source by lymphocytes (Jenkins et al., 1999). In our study, supplemental dextrin increased fecal SCFA concentrations. Most notable was the increase in butyrate production, which has been associated with reductions in inflammatory genes, increased growth performance, increased intestinal absorptive capacity, and decreased *E. coli* abundance (Lu et al., 2008).

In the current study, supplemental dextrin altered the composition of the gut microbiome, as evidenced by a statistical difference in the Bray Curtis index. Supplementing dextrin during the pre-weaning period caused reduction in some bacteria associated with gut inflammation, in addition to an increase in bacteria associated with beneficial functions. Pigs that received dextrin during the pre-weaning period (F/F and F/NF groups) had an increase in some SCFA producing bacteria ($q < 0.05$) compared to pigs that never received dextrin (NF/NF). Pigs fed dextrin during the pre-weaning period had an increase in beneficial bacteria, including the genera *Coprococcus, Butyricimonas, Anaerovibrio, Paraprevotella,* and *Clostridium sensu stricto,* which have previously been shown to produce SCFAs (Polansky et al., 2016; Sakamoto et al., 2009; Vital et al., 2014). Thus, these bacteria may produce SCFAs in the pig hindgut, promote healthy immune pathways, and maintain gut homestasis (Bressa et al., 2017; Lopetuso et al., 2013; Sun et al., 2016; F. Yang et al., 2020). Other OTUs were enriched, but their effect on gut health remains unclear. For example, *Mogibacterium* has been found to be a core member of the gut microbiota but not a SCFA producer (L. Chen et al., 2017) while *Bilophila* has been associated with low performing swine (Gardiner et al., 2020), but has not been shown to cause disease in the intestines (McOrist et al., 2001). In both cecum and colon of pigs that never received dextrin (NF/NF) there was an increased relative abundance of unclassified *Desulfovibrionaceae* ($q < 0.05$), a family with members implicated in increased inflammation (Figliuolo et al., 2017). Pigs that received dextrin only in the post-weaning period (NF/F) seem to be of an intermediate community type. There were no genera with different relative abundance with the NF/NF group and only one genera different with the F/F group. Thus, dextrin supplementation in the post-weaning period did not result in the microbial community shift like pre-weaning dextrin supplementation.

In addition to inflammatory markers increasing, weaning is associated with reductions in villus height and increases in crypt depths (Campbell et al., 2013). This is indicative of intestinal inflammation as well as increased intestinal cell sloughing (Land, 2015). In this study, no changes

in villi heights were observed between treatment groups. However, a trend for a reduction in crypt depths was observed with dextrin supplementation, leading to a trend for an increase in villus:crypt. This implies that addition of dextrin did not increase the absorptive capacity in the ileum, but potentially reduced the amount of stress placed on the small intestine (Nabuurs et al., 1993). Seeing this reduction in intestinal stress coupled with SCFA production, growth performance would be expected to increase. The only statistically significant difference in growth performance data showed pigs fed dextrin prior to weaning were more efficient from day 11 until day 21. However, when analyzed over the entire study, there were no differences in ADG or ADFI among diet groups. **W**hile prebiotic changes benefit the health of the gut tissue and immune response (as discussed above), animal performance is not always increased (Jha et al., 2019; Markowiak & Śliżewska, 2018; Ngoc et al., 2013)**.** In one study, diet supplementation with a mixture of galactooligosaccharides did not alter animal growth but decreased attachment of enterohepatic *Escherichia coli* and *Salmonella enterica* serotype Typhimurium (Tzortzis et al., 2005). This suggests that while animals fed supplemental prebiotics may not have increased growth rate, due to altered immune status and intestinal tissue health they might be able to perform better under a pathogen challenge or other suboptimal conditions,.

Pigs that received supplemental dextrin only prior to weaning or only post-weaning (F/NF or NF/F, respectively) typically showed a more beneficial response than the pigs that received supplemental dextrin for the entirety of the study (F/F). This was seen for response variables such as total SCFA concentrations, acetate concentration, and butyrate concentration. These results were surprising as we expected pigs receiving supplemental dextrin both pre- and post-weaning to have an increased response to these variables. This could be due to differences in carbohydrase production. Prior to weaning, lactase is the main carbohydrase being produced to digest the lactose in the sow's milk (Hartman et al., 1961). With the addition of dextrin, different carbohydrases will be produced since dextrin is capable of being broken down in the small intestine due to the α-1,4 linkage (Singh et al., 2010; Takata et al., 2005). This could prepare the pig to better digest the multitude of carbohydrates that are available in solid feed after weaning. Pigs fed supplemental dextrin prior to weaning could then be expected to perform better than those that had not received supplemental dextrin prior to weaning. Pigs fed supplemental dextrin for the entirety of the study then would digest more of the dextrin fed post-weaning, which would limit the amount of dextrin available to the cecum to be fermented. Pigs being fed dextrin only post-weaning would be

expected to not have as high carbohydrase production to metabolizing dextrin in the small intestine, allowing more dextrin to make it to the cecum and be fermented by the bacteria to produce more SCFAs. The change in carbohydrase production could explain the increases in the response variables that we observed for pigs that only received supplemental dextrin either pre- or post-weaning.

From the microbiome perspective, it appears that supplemental dextrin fed only during the post-weaning period (NF/F) resulted in the fewest changes in bacterial taxa compared to the other diet groups, while pre-weaning dextrin (F/NF) appears to have promoted many beneficial bacteria compared with NF/NF. Additionally, when animals were fed dextrin during the pre-weaning period but differed in the post-weaning period (F/F and F/NF groups), there were still many OTUs that had different relative abundance. Interestingly, many of the assumed functions of the differentially abundant OTUs between these two groups were either increased or decreased in both diet groups. For example, *Butyricimonas* and *Coprococcus* both produce beneficial SCFAs (Nogal et al., 2021; Sakamoto et al., 2009) but *Butyricimonas* was increased in the F/NF group, while *Coprococcus* was increased in the F/F group. This suggests that supplementing dextrin post-weaning may not have changed the function of the community as a whole, but instead shifted which populations present to carry out these functions. This was previously seen in the human gut – higher order taxa of microbial communities do not shift due to small or short-term diet changes; the genera may change but the functions they provide are similar (Arumugam et al., 2011). Thus, it appears that some community divergence occurred during the post-weaning period, but the overall community function appears similar between the two groups.

Mucosal communities were significantly different from digesta communities in piglets after weaning, as has been determined previously (Burrough et al., 2017). In addition, dextrin supplementation impacted the mucosal bacterial community differently than the luminal bacterial community in the current study, as is seen in the types of bacteria with increased abundance in the mucosal samples - e.g. *Muscispirillum*, a bacteria commonly associated with diabetes and colonic inflammation (Berry et al., 2012; Robertson et al., 2005). The mucosal-associated bacterial communities may be of increased importance to the gut health of animals (Aidy et al., 2013; Awad et al., 2016; Van den Abbeele et al., 2013). While metabolites produced in the lumen can still have an impact on the animal (Donaldson et al., 2016; Kong et al., 2016), mucosal microbial communities are thought to be of more relevance to gut health because mucosal communities have

a longer retention time in the gut, are in closer contact to the epithelium, and thus can have a more direct impact on the structure and function of the intestinal barrier (Finnie et al., 1995; Martens et al., 2008; Van den Abbeele et al., 2013). Thus, these shifts to the mucosal bacterial community may play an important role in the gut health of dextrin supplemented pigs.

3.5 Conclusion

Data from this study indicates that feeding supplemental soluble fiber prior to and/or after weaning resulted in changes in SCFA production, crypt depth, and the cecal and colon microbiome. There were some benefits from feeding dextrin to pigs around weaning, e.g., increase in SCFA production. To increase SCFA production, it may be beneficial to supplement soluble fiber to pigs post-weaning. Feeding piglets dextrin prior to weaning resulted in potentially beneficial modulations of the microbiome but a practical method to feed soluble fiber before weaning would need to be determined. Though some beneficial effects from the addition of supplemental dextrin were observed, the effects tended to be smaller than expected. Different types and sources of fiber have been reported to produce different effects in pigs, and additional investigation is warranted.

3.6 References

Aidy, S. E., Abbeele, P. V. den, Wiele, T. V. de, Louis, P., & Kleerebezem, M. (2013). Intestinal colonization: How key microbial players become established in this dynamic process. BioEssays, 35(10), 913–923. https://doi.org/10.1002/bies.201300073

Allen, H. K. (2017). Alternatives to antibiotics: why and how. NAM Perspectives, 7(7). https://doi.org/10.31478/201707g

Allen, H. K., Levine, U. Y., Looft, T., Bandrick, M., & Casey, T. A. (2013). Treatment, promotion, commotion: Antibiotic alternatives in food-producing animals. Trends in Microbiology, 21(3), 114–119. https://doi.org/10.1016/j.tim.2012.11.001

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., … Bork, P. (2011). Enterotypes of the human gut microbiome. Nature, 473(7346), 174–180. https://doi.org/10.1038/nature09944

Awad, W. A., Mann, E., Dzieciol, M., Hess, C., Schmitz-Esser, S., Wagner, M., & Hess, M. (2016). Age-related differences in the luminal and mucosa-associated gut microbiome of broiler chickens and shifts associated with Campylobacter jejuni infection. Frontiers in Cellular and Infection Microbiology, 6, 223–17. https://doi.org/10.3389/fcimb.2016.00154

Bassaganya-Riera, J., DiGuardo, M., Viladomiu, M., de Horna, A., Sanchez, S., Einerhand, A. W. C., Sanders, L., & Hontecillas, R. (2011). Soluble fibers and resistant starch ameliorate disease activity in interleukin-10–deficient mice with inflammatory bowel disease. The Journal of Nutrition, 141(7), 1318–1325. https://doi.org/10.3945/jn.111.139022

Bauer, E., Williams, B. A., Verstegen, M. W. A., & Mosenthin, R. (2006). Chapter 2 Fermentable carbohydrates: Potential dietary modulators of intestinal physiology, microbiology and immunity in pigs. In Biology of Growing Animals (Vol. 4, pp. 33–63). Elsevier. https://doi.org/10.1016/S1877-1823(09)70089-X

Berry, D., Schwab, C., Milinovich, G., Reichert, J., Ben Mahfoudh, K., Decker, T., Engel, M., Hai, B., Hainzl, E., Heider, S., Kenner, L., Müller, M., Rauch, I., Strobl, B., Wagner, M., Schleper, C., Urich, T., & Loy, A. (2012). Phylotype-level 16S rRNA analysis reveals new bacterial indicators of health state in acute murine colitis. The ISME Journal, 6(11), 2091–2106. https://doi.org/10.1038/ismej.2012.39

Bressa, C., Bailén-Andrino, M., Pérez-Santiago, J., González-Soltero, R., Pérez, M., Montalvo-Lominchar, M. G., Maté-Muñoz, J. L., Domínguez, R., Moreno, D., & Larrosa, M. (2017). Differences in gut microbiota profile between women with active lifestyle and sedentary women. PLOS ONE, 12(2), e0171352. https://doi.org/10.1371/journal.pone.0171352

Burrough, E. R., Arruda, B. L., & Plummer, P. J. (2017). Comparison of the luminal and mucosa-associated microbiota in the colon of pigs with and without swine dysentery. Frontiers in Veterinary Science, 4, 139. https://doi.org/10.3389/fvets.2017.00139

Calbrix, R., Guernsey, J., Schinckel, A., Stewart, T., Herman, E., Helm, R., & Nielsen, N. (2012). Development of tools to study immune-mediated allergenic responses to food and feed. In Designing Soybeans for 21st Century Markets (pp. 239–252). Elsevier. https://doi.org/10.1016/B978-0-9830791-0-1.50017-0

Campbell, J. M., Crenshaw, J. D., & Polo, J. (2013). The biological stress of early weaned piglets. Journal of Animal Science and Biotechnology, 4(1), 19. https://doi.org/10.1186/2049-1891-4-19

Chen, H., Mao, X., He, J., Yu, B., Huang, Z., Yu, J., Zheng, P., & Chen, D. (2013). Dietary fibre affects intestinal mucosal barrier function and regulates intestinal bacteria in weaning piglets. British Journal of Nutrition, 110(10), 1837–1848. https://doi.org/10.1017/S0007114513001293

Chen, H., Wang, C., Wang, Y., Chen, Y., Wan, M., Zhu, J., & Zhu, A. (2021). Effects of soft pellet creep feed on pre-weaning and post-weaning performance and intestinal development in piglets. Animal Bioscience, 34(4), 714–723. https://doi.org/10.5713/ajas.20.0034

Chen, L., Xu, Y., Chen, X., Fang, C., Zhao, L., & Chen, F. (2017). The maturing development of gut microbiota in commercial piglets during the weaning transition. Frontiers in Microbiology, 8, 1688. https://doi.org/10.3389/fmicb.2017.01688

Choudhury, R., Middelkoop, A., Boekhorst, J., Gerrits, W. J. J., Kemp, B., Bolhuis, J. E., & Kleerebezem, M. (2021). Early life feeding accelerates gut microbiome maturation and suppresses acute post-weaning stress in piglets. Environmental Microbiology, 23(11), 7201–7213. https://doi.org/10.1111/1462-2920.15791

Christensen, B., & Huber, L.-A. (2021). The effect of creep feed composition and form on pre- and post-weaning growth performance of pigs and the utilization of low-complexity nursery diets. Translational Animal Science, 5(4), txab211. https://doi.org/10.1093/tas/txab211

Cleave, T. L. (1968). Sucrose intake and coronary heart-disease. Originally Published as Volume 2, Issue 7579, 292(7579), 1187. https://doi.org/10.1016/S0140-6736(68)91659-0

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., & Tiedje, J. M. (2013). Ribosomal Database Project: Data and tools for high throughput rRNA analysis. Nucleic Acids Research, 42(D1), D633–D642. https://doi.org/10.1093/nar/gkt1244

Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. Nature Reviews Microbiology, 14(1), 20–32. https://doi.org/10.1038/nrmicro3552

Figliuolo, V. R., dos Santos, L. M., Abalo, A., Nanini, H., Santos, A., Brittes, N. M., Bernardazzi, C., de Souza, H. S. P., Vieira, L. Q., Coutinho-Silva, R., & Coutinho, C. M. L. M. (2017). Sulfate-reducing bacteria stimulate gut immune responses and contribute to inflammation in experimental colitis. Life Sciences, 189, 29–38. https://doi.org/10.1016/j.lfs.2017.09.014

Finnie, I. A., Dwarakanath, A. D., Taylor, B. A., & Rhodes, J. M. (1995). Colonic mucin synthesis is increased by sodium butyrate. Gut, 36(1), 93–99. https://doi.org/10.1136/gut.36.1.93

Franklin, M. A., Mathew, A. G., Vickers, J. R., & Clift, R. A. (2002). Characterization of microbial populations and volatile fatty acid concentrations in the jejunum, ileum, and cecum of pigs weaned at 17 vs 24 days of age. Journal of Animal Science, 80(11), 2904–2910. https://doi.org/10.2527/2002.80112904x

Frese, S. A., Parker, K., Calvert, C. C., & Mills, D. A. (2015). Diet shapes the gut microbiome of pigs during nursing and weaning. Microbiome, 3(1), 28. https://doi.org/10.1186/s40168-015-0091-8

Gardiner, G. E., Metzler-Zebeli, B. U., & Lawlor, P. G. (2020). Impact of intestinal microbiota on growth and feed efficiency in pigs: A review. Microorganisms, 8(12), 1886. https://doi.org/10.3390/microorganisms8121886

Guevarra, R. B., Hong, S. H., Cho, J. H., Kim, B.-R., Shin, J., Lee, J. H., Kang, B. N., Kim, Y. H., Wattanaphansak, S., Isaacson, R. E., Song, M., & Kim, H. B. (2018). The dynamics of the piglet gut microbiome during the weaning transition in association with health and nutrition. Journal of Animal Science and Biotechnology, 9(1), 54. https://doi.org/10.1186/s40104-018-0269-6

Hartman, P. A., Hays, V. W., Baker, R. O., Neagle, L. H., & Catron, D. V. (1961). Digestive enzyme development in the young pig. Journal of Animal Science, 20(1), 114–123. https://doi.org/10.2527/jas1961.201114x

Hötzel, M. J., de Souza, G. P. P., Costa, O. A. D., & Machado Filho, L. C. P. (2011). Disentangling the effects of weaning stressors on piglets' behaviour and feed intake: Changing the housing and social environment. Applied Animal Behaviour Science, 135(1–2), 44–50. https://doi.org/10.1016/j.applanim.2011.09.003

Jarrett, S., & Ashworth, C. J. (2018). The role of dietary fibre in pig production, with a particular emphasis on reproduction. Journal of Animal Science and Biotechnology, 9(1), 59. https://doi.org/10.1186/s40104-018-0270-0

Jenkins, D. J. A., Kendall, C. W. C., Augustin, L. S. A., Martini, M. C., Axelsen, M., Faulkner, D., Vidgen, E., Parker, T., Lau, H., Connelly, P. W., Teitel, J., Singer, W., Vandenbroucke, A. C., Leiter, L. A., & Josse, R. G. (2002). Effect of wheat bran on glycemic control and risk factors for cardiovascular disease in type 2 diabetes. Diabetes Care, 25(9), 1522–1528. https://doi.org/10.2337/diacare.25.9.1522

Jenkins, D. J. A., Kendall, C. W. C., & Vuksan, V. (1999). Inulin, oligofructose and intestinal function. The Journal of Nutrition, 129(7), 1431S-1433S. https://doi.org/10.1093/jn/129.7.1431S

Jha, R., Fouhse, J. M., Tiwari, U. P., Li, L., & Willing, B. P. (2019). Dietary Fiber and Intestinal Health of Monogastric Animals. Frontiers in Veterinary Science, 6, 48. https://doi.org/10.3389/fvets.2019.00048

Knapp, B., Bauer, L., Swanson, K., Tappenden, K., Fahey, G., & de Godoy, M. (2013). Soluble fiber dextrin and soluble corn fiber supplementation modify indices of health in cecum and colon of sprague-dawley rats. Nutrients, 5(2), 396–410. https://doi.org/10.3390/nu5020396

Kong, X., Ji, Y., Li, H., Zhu, Q., Blachier, F., Geng, M., Chen, W., & Yin, Y. (2016). Colonic luminal microbiota and bacterial metabolite composition in pregnant Huanjiang mini-pigs: Effects of food composition at different times of pregnancy. Scientific Reports, 6(1), 37224. https://doi.org/10.1038/srep37224

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

Lallès, J.-P., Bosi, P., Smidt, H., & Stokes, C. R. (2007). Weaning—A challenge to gut physiologists. Livestock Science, 108(1–3), 82–93. https://doi.org/10.1016/j.livsci.2007.01.091

Land, W. G. (2015). The role of damage-associated molecular patterns (DAMPs) in human diseases poart II: DAMPs as diagnostics, prognostics and therapeutics in clinical medicine. Sultan Qaboos University Medical Journal, 15(2), 157–170.

Lee, S. A., Febery, E., Wilcock, P., & Bedford, M. R. (2021). Application of creep feed and phytase super-dosing as tools to support digestive adaption and feed efficiency in piglets at weaning. Animals, 11(7), 2080. https://doi.org/10.3390/ani11072080

Lopetuso, L. R., Scaldaferri, F., Petito, V., & Gasbarrini, A. (2013). Commensal Clostridia: Leading players in the maintenance of gut homeostasis. Gut Pathogens, 5(1), 23. https://doi.org/10.1186/1757-4749-5-23

Lu, J., Zou, X., & Wang, Y. (2008). Effects of sodium butyrate on the growth performance, intestinal microflora and morphology of weanling pigs. Journal of Animal and Feed Sciences, 17(4), 568–578. https://doi.org/10.22358/jafs/66685/2008

Markowiak, P., & Śliżewska, K. (2018). The role of probiotics, prebiotics and synbiotics in animal nutrition. Gut Pathogens, 10(1), 21. https://doi.org/10.1186/s13099-018-0250-0

Martens, E. C., Chiang, H. C., & Gordon, J. I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host & Microbe, 4(5), 447–457. https://doi.org/10.1016/j.chom.2008.09.007

McKAY, D. M., & Baird, A. W. (1999). Cytokine regulation of epithelial permeability and ion transport. Gut, 44(2), 283–289. https://doi.org/10.1136/gut.44.2.283

McOrist, A. L., Warhurst, M., McOrist, S., & Bird, A. R. (2001). Colonic infection by *Bilophila wadsworthia* in pigs. Journal of Clinical Microbiology, 39(4), 1577–1579. https://doi.org/10.1128/JCM.39.4.1577-1579.2001

Middelkoop, A., Choudhury, R., Gerrits, W. J. J., Kemp, B., Kleerebezem, M., & Bolhuis, J. E. (2020). Effects of creep feed provision on behavior and performance of piglets around weaning. Frontiers in Veterinary Science, 7, 520035. https://doi.org/10.3389/fvets.2020.520035

Moeser, A. J., Klok, C. V., Ryan, K. A., Wooten, J. G., Little, D., Cook, V. L., & Blikslager, A. T. (2007). Stress signaling pathways activated by weaning mediate intestinal dysfunction in the pig. American Journal of Physiology-Gastrointestinal and Liver Physiology, 292(1), G173–G181. https://doi.org/10.1152/ajpgi.00197.2006

Nabuurs, M. J. A., Hoogendoorn, A., Van Der Molen, E. J., & Van Osta, A. L. M. (1993). Villus height and crypt depth in weaned and unweaned pigs, reared under various circumstances in the Netherlands. Research in Veterinary Science, 55(1), 78–84. https://doi.org/10.1016/0034-5288(93)90038-H

Ngoc, T. T. B., Len, N. T., & Lindberg, J. E. (2013). Impact of fibre intake and fibre source on digestibility, gut development, retention time and growth performance of indigenous and exotic pigs. Animal, 7(5), 736–745. https://doi.org/10.1017/S1751731112002169

Nogal, A., Louca, P., Zhang, X., Wells, P. M., Steves, C. J., Spector, T. D., Falchi, M., Valdes, A. M., & Menni, C. (2021). Circulating levels of the short-chain fatty acid acetate mediate the effect of the gut microbiome on visceral fat. Frontiers in Microbiology, 12. https://www.frontiersin.org/article/10.3389/fmicb.2021.711359

Oswald, I. P. (2006). Role of intestinal epithelial cells in the innate immune defence of the pig intestine. Veterinary Research, 37(3), 359–368. https://doi.org/10.1051/vetres:2006006

Pluske, J. R., Hampson, D. J., & Williams, I. H. (1997). Factors influencing the structure and function of the small intestine in the weaned pig: A review. Livestock Production Science, 51(1–3), 215–236. https://doi.org/10.1016/S0301-6226(97)00057-2

Pluske, J. R., Williams, I. H., & Aherne, F. X. (1996). Maintenance of villous height and crypt depth in piglets by providing continuous nutrition after weaning. Animal Science, 62(1), 131–144. https://doi.org/10.1017/S1357729800014417

Polansky, O., Sekelova, Z., Faldynova, M., Sebkova, A., Sisak, F., & Rychlik, I. (2016). Important metabolic pathways and biological processes expressed by chicken cecal microbiota. Applied and Environmental Microbiology, 82(5), 1569–1576. https://doi.org/10.1128/AEM.03473-15

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glockner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Research, 41(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/

Rhouma, M., Fairbrother, J. M., Beaudry, F., & Letellier, A. (2017). Post weaning diarrhea in pigs: Risk factors and non-colistin-based control strategies. Acta Veterinaria Scandinavica, 59(1), 31. https://doi.org/10.1186/s13028-017-0299-7

Robertson, B. R., O'Rourke, J. L., Neilan, B. A., Vandamme, P., On, S. L. W., Fox, J. G., & Lee, A. (2005). Mucispirillum schaedleri gen. Nov., sp. Nov., a spiral-shaped bacterium colonizing the mucus layer of the gastrointestinal tract of laboratory rodents. International Journal of Systematic and Evolutionary Microbiology, 55(3), 1199–1204. https://doi.org/10.1099/ijs.0.63472-0

Sakamoto, M., Takagaki, A., Matsumoto, K., Kato, Y., Goto, K., & Benno, Y. (2009). Butyricimonas synergistica gen. Nov., sp. Nov. and Butyricimonas virosa sp. Nov., butyric acid-producing bacteria in the family "Porphyromonadaceae" isolated from rat faeces. INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY, 59(7), 1748–1753. https://doi.org/10.1099/ijs.0.007674-0

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., & Weber, C. F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology, 75(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09

Shim, S. B., Verstegen, M. W. A., Kim, I. H., Kwon, O. S., & Verdonk, J. M. A. J. (2005). Effects of feeding antibiotic-free creep feed supplemented with oligofructose, probiotics or synbiotics to suckling piglets increases the preweaning weight gain and composition of intestinal microbiota. Archives of Animal Nutrition, 59(6), 419–427. https://doi.org/10.1080/17450390500353234

Singer, R. S., Porter, L. J., Thomson, D. U., Gage, M., Beaudoin, A., & Wishnie, J. K. (2019). Raising animals without antibiotics: U.S. producer and veterinarian experiences and opinions. Frontiers in Veterinary Science, 6, 452. https://doi.org/10.3389/fvets.2019.00452

Singh, J., Dartois, A., & Kaur, L. (2010). Starch digestibility in food matrix: A review. Trends in Food Science & Technology, 21(4), 168–180. https://doi.org/10.1016/j.tifs.2009.12.001

Sun, Y., Su, Y., & Zhu, W. (2016). Microbiome-metabolome responses in the cecum and colon of pig to a high resistant starch diet. Frontiers in Microbiology, 7. https://doi.org/10.3389/fmicb.2016.00779

Takata, H., Kojima, I., Taji, N., Suzuki, Y., & Yamamoto, M. (2005). Industrial production of branching enzyme, and its application for producing a highly branched cyclic dextrin, Cluster DextrinTM. Proceeding of Current Biotechnology and Bioengineering, 2005.10, 119–120. http://203.250.217.22/article/CFKO200536037085934.pdf

Trowell, H. (1972). Ischemic heart disease and dietary fiber. The American Journal of Clinical Nutrition, 25(9), 926–932. https://doi.org/10.1093/ajcn/25.9.926

Turner, J. R. (2006). Molecular basis of epithelial barrier regulation. The American Journal of Pathology, 169(6), 1901–1909. https://doi.org/10.2353/ajpath.2006.060681

Tzortzis, G., Goulas, A. K., Gee, J. M., & Gibson, G. R. (2005). A novel galactooligosaccharide mixture increases the bifidobacterial population numbers in a continuous in vitro fermentation system and in the proximal colonic contents of pigs in vivo. The Journal of Nutrition, 135(7), 1726–1731. https://doi.org/10.1093/jn/135.7.1726

Van den Abbeele, P., Belzer, C., Goossens, M., Kleerebezem, M., De Vos, W. M., Thas, O., De Weirdt, R., Kerckhof, F.-M., & Van de Wiele, T. (2013). Butyrate-producing Clostridium cluster XIVa species specifically colonize mucins in an in vitro gut model. The ISME Journal, 7(5), 949–961. https://doi.org/10.1038/ismej.2012.158

Vital, M., Howe, A. C., & Tiedje, J. M. (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. MBio, 5(2). https://doi.org/10.1128/mBio.00889-14

Whisner, C. M., Martin, B. R., Nakatsu, C. H., Story, J. A., MacDonald-Clarke, C. J., McCabe, L. D., McCabe, G. P., & Weaver, C. M. (2016). Soluble corn fiber increases calcium absorption associated with shifts in the gut microbiome: A randomized dose-response trial in free-living pubertal females. The Journal of Nutrition, 146(7), 1298–1306. https://doi.org/10.3945/jn.115.227256

Wijtten, P. J. A., Meulen, J. van der, & Verstegen, M. W. A. (2011). Intestinal barrier function and absorption in pigs after weaning: A review. British Journal of Nutrition, 105(7), 967–981. https://doi.org/10.1017/S0007114510005660

Yang, F., Zhang, S., Tian, M., Chen, J., Chen, F., & Guan, W. (2020). Different sources of high fat diet induces marked changes in gut microbiota of nursery pigs. Frontiers in Microbiology, 11, 859. https://doi.org/10.3389/fmicb.2020.00859

PART II – BACTERIAL POPULATION ANALYSIS

CHAPTER 4. ANTIBIOTIC RESISTANCE DETERMINATION USING MACHINE LEARNING IMPROVES GENOTYPE-PHENOTYPE CONCORDANCE IN BACTERIAL PATHOGENS ASSOCIATED WITH BOVINE RESPIRATORY DISEASE

Carmen L. Wickware[1], Mohit Verma[2], and Timothy A. Johnson[1]
Purdue University, [1]Department of Animal Sciences and [2]Department of Agricultural and Biological Engineering, West Lafayette, IN 47907 USA

Acknowledgement: This work was supported by the Agriculture and Food Research Initiative Competitive Grants Program Award 2020-68014-31302 from the U.S. Department of Agriculture National Institute of Food and Agriculture. This work was also supported by a Purdue University Agricultural Science and Extension for Economic Development (AgSEED) grant. The funding organizations played no role in the design of the study; collection, analysis, and interpretation of data; or in writing the manuscript.

Abstract

Bovine respiratory disease (BRD) is one of the costliest diseases for producers, along with having a high rate of mortality for both dairy and beef cattle. Since microbiological testing takes longer than can be afforded to prevent morbidity and mortality, antibiotic therapy is often administered empirically. Phenotypic antibiotic resistance is common in BRD pathogens and rapid genetic assays to identify antibiotic resistance genes in pathogenic bacteria could be useful for determining antibiotic resistance phenotypes. However, when relying on previously characterized antibiotic resistance genes, genotype-phenotype concordance rates are low in BRD pathogens compared to well-studied pathogens such as *Salmonella* Typhimurium. The objective of this study was to use a machine learning approach to determine marker sequences of antibiotic resistance in three BRD pathogens, *Mannheimia haemolytica*, *Pasteurella multocida*, and *Histophilus somni*. Previously established antibiotic resistance genotype-phenotype concordance for these pathogens is low for many tested antibiotics, suggesting there are unknown genes or sequences determining resistance. **We hypothesized that the use of machine learning methods would improve our ability to predict phenotypic antibiotic resistance by lowering the number of false negative predictions.** The error rates for predicting phenotypic resistance to tulathromycin and tilmicosin, two antibiotics used to treat BRD, using known antibiotic resistance genes alone were 41% and 39%, respectively. By using marker sequences generated with machine learning, the error rates

were decreased to 7.4% and 15% for tulathromycin and tilmicosin, respectively. A similar improvement of error rate in the predicted phenotype was seen in all tested antibiotics. The marker sequences found for the tested antibiotics were annotated as transposases, tetracycline repressor *tetR*, and some amino acid biosynthesis genes. Regions surrounding the resistance marker sequences (i.e., found on the same contig) contained other known resistance genes (*sul2, bla*$_{ROB}$, *arsR*) and some had conjugation machinery genes. However, the identified marker genes of antibiotic resistance were not annotated as known antibiotic resistance genes, but were annotated as CRISPR arrays, histidine biosynthesis genes, transposases, and mobile element proteins. Mobile genetic elements (MGEs) such as transposases can be associated with many different resistance genes. Despite the individual implications of the marker sequences, using machine learning to generate markers of antibiotic resistance increased the chances of accurately determining the resistance phenotype of the bacterial isolates. With the marker sequences from this study, assays could be developed to rapidly determine resistance phenotypes, which could help improve treatment success in cattle with BRD.

4.1 Introduction

Bovine respiratory disease (BRD) is a broad term for diseases affecting the respiratory tract in cattle. With approximately 20-30% of a herd affected per year, BRD is one of the costliest diseases for both beef and dairy cattle producers (Peel, 2020; Theurer et al., 2021). While the estimated economic loss from BRD is $800-900 million annually from treatment, retreatment, low feed efficiency, and replacement animals (Peel, 2020; Theurer et al., 2021), BRD also results in increase loss of animal lives and welfare. Many factors increase the risk of an animal becoming susceptible to BRD including environmental changes and management practices (e.g., temperature fluctuations and stocking density).

However, these factors are not the only concerns in the complexity of BRD as epidemiological risk factors are the direct cause of disease. BRD is often polymicrobial, meaning it can result from multiple microbes (e.g., a viral infection followed by a bacterial infection) (Agnes et al., 2013; Petruzzi et al., 2020). Many of the bacterial pathogens associated with BRD are opportunistic; they exist as a part of the commensal community and become pathogenic when the conditions allow them to by-pass animal immune defenses and/or outgrow other commensal bacteria (Murray et al., 2016; Sandal & Inzana, 2010).

Detection of BRD through visual observation of animals combined with clinical measurements (the current established method) has low accuracy, but waiting for a laboratory diagnosis can take several days or longer (White & Renter, 2009). As an alternative, a rapid biosensor, such as has been used in the diagnosis of Sars-COV-2 (Davidson et al., 2021), could be used to determine the most likely causative agent of BRD (Pascual-Garrigos et al., 2021). Rapid determination of the causative agent of BRD and the antibiotic resistance profile of the pathogen may provide the data necessary to determine the most appropriate antibiotic(s) for treating BRD or if antibiotics should not be administered.

Targeting specific antibiotic resistance genes with polymerase chain reaction (PCR) determine resistant bacteria, a common method, only produces output on if the bacteria have the capacity to be resistant to antibiotics, not if they are physiologically resistant. Previous methods to determine novel resistance genes or predict resistance have relied on comparison of the genome in question to reference sequences and genomes to identify novel antibiotic resistance genes based on homology to known genes (Knopp et al., 2021; Köser et al., 2012; Wood et al., 2020) as well as mining assembled and unassembled metagenomes for sequences with similarity to known resistance genes (Berglund et al., 2019; Pehrsson et al., 2013; Willms et al., 2019). However, these methods only detect sequences with similarity to previously identified resistance genes, not if the presence of the genes confer resistance to the bacterium that encodes the genes.

Machine learning allows the use of genome sequences in addition to antibiotic resistance phenotypes from a large training set of isolates as input and determines *de novo* markers of antibiotic resistance without relying on currently available resistance gene databases. Using machine learning to identify marker sequences of phenotypic antibiotic resistance may improve the prediction of resistance phenotypes from genomic sequences and potentially identify novel resistance genes for further characterization. For this study, we used a machine learning program, Kover, to predict phenotypic antibiotic resistance from genome assembly data without the use of ARG reference sequences (Drouin et al., 2016, 2019).

4.2 Materials and Methods

4.2.1 Isolates

Isolates of *Mannheimia haemolytica* (n=6)*, Pasteurella multocida* (n=7)*,* and *Histophilus somni* (n=5) were obtained from cattle nasopharyngeal and deep lung swabs submitted to the Purdue University Animal Disease Diagnostic Laboratory (PA). PA isolates were prepared for genome sequencing using the TruSeq DNA PCR-free Library Preparation kit (Illumina, San Diego, USA). Sequencing was performed at the Purdue Genomics Core using MiSeq sequencing (Illumina; 2 x 300 cycles).

4.2.2 Genome Sequences and Assembly

Isolates from BRD pathogen species – *M. haemolytica* (n=26), *P. multocida* (n=25), and *H. somni* (n=13) – previously sequenced and assembled by Owen et al (OA; 2017) were also used in the current study together with the PA isolates. OA assembled genomes were downloaded from the NCBI Genome database (Bioproject PRJNA306895). PA genome sequences were assembled using SPAdes (Bankevich et al., 2012; Nurk et al., 2013) with default parameters for assembly but including the *"--careful"* parameter to reduce the number of mismatches and short contigs.

Assembly quality was assessed using quast v3.2 (Gurevich et al., 2013; Mikheenko et al., 2016) with default parameters. Recorded statistics include number of contigs, max contig, total length, and N50. Full assembly statistics for PA isolates are listed in Appendix B.1.

4.2.3 Antibiotic susceptibility testing (AST) and ARG annotation

Antibiotic susceptibility of OA isolates was tested using Sensititre Bovine/Porcine AST plate BOPO6F (Thermo Scientific; Owen et al., 2017) while PA isolates were tested using Bovine AST plate BOPO7F (Thermo Scientific). Some antibiotics (tildipirosin and those related to tetracycline) are not included on both plates. In these cases, antibiotics for which a minimum inhibitory concentration (MIC) was not determined for all isolates were removed to produce a balanced dataset. MIC values and interpretations for PA isolates are listed in Appendix B.2 and are listed in supplementary table 1 (TableS1) of Owen et al., (2017) for OA isolates.

ARGs were annotated using the Complete Antibiotic Resistance Database - Resistance Gene Identifier (CARD-RGI, Alcock et al., 2019). Though Owen et al annotated ARGs in their study, , we re-annotated the OA genomes with CARD-RGI to limit database bias between the two sets of isolates. Parameters used for CARD-RGI were as follows:

--input_type contig -d wgs --local --exclude_nudge –clean

The "loose" matches were excluded, leaving only "strict" and "perfect" annotations, the highest confidence gene calls from CARD-RGI. Concordance (accuracy) was calculated for each antibiotic by taking the sum of the isolates with matching genotype (G) and phenotype (P) of antibiotic susceptible (AS) or resistant (AR) isolates (G-AR+P-AR and G-AS+P-AS) and dividing by the total sum of isolates, i.e., G-AR/P-AR and G-AS/P-AR and G-AR/P-AS and G-AS/P-AS (Table 4.1). Additional metrics sensitivity (recall), specificity, precision, and F1 score were also calculated for each antibiotic.

4.2.4 Antibiotic resistance prediction using machine learning

Assembled genomes were used to create a table of k-mers (k=31) representing the presence or absence of sequences 31 bases in length in each genome. K-mer length 31 was used as it produced the best models (data not shown). The k-mers were created using Suffixerator and counted using Tallymer through the GenomeTools wrapper program (v 1.5.9; Kurtz et al., 2008; Gremme et al., 2013). A metadata file was created for the isolates, which contained the assembly names and whether the isolates were phenotypically resistant (1) or susceptible (0) to a particular antibiotic. A separate prediction model was created for each antibiotic. To create a model of AR prediction, more than 10 isolates in both the resistant and susceptible groups were required (Drouin et al., 2016). The k-mer table and metadata were both used as input files to predict resistance in Kover with the Set Covering Machine algorithm (Drouin et al., 2016, 2019; Marchand & Shawe-Taylor, 2002). Briefly, the dataset was split into training and testing groups (66 and 33% of the data, respectively). A model was learned using the training data and validated using the testing data. The output was a set of rules (31-mers) each assigned an importance based on how often the rule was considered for the model (Drouin et al., 2016, 2019). Also included as output were the equivalent rules (ER) – the set of 31-mers considered equally important for predicting AR in the model.

Table 4.1 Concordance between antibiotic resistance phenotype and genes. P: phenotype; G: genotype; R: resistant; S: susceptible; ML: machine learning

	P:R		P:S						
ML Antibiotics	G:R	G:S	G:R	G:S	Precision	Recall/Sensitivity	Specificity	F1 score	Accuracy
Ampicillin	17	4	7	54	0.71	0.81	0.89	0.76	0.87
Clindamycin	42	36	4	0	0.91	0.54	0.00	0.68	0.51
Danofloxacin	0	27	0	55	N/A	0.00	1.00	N/A	0.67
Penicillin	21	5	3	53	0.88	0.81	0.95	0.84	0.90
Tiamulin	0	25	0	57	N/A	0.00	1.00	N/A	0.70
Tilmicosin	34	20	12	16	0.74	0.63	0.57	0.68	0.61
Tulathromycin	12	0	34	36	0.26	1.00	0.51	0.41	0.59
	P:R		P:S						
Non-ML antibiotics	G:R	G:S	G:R	G:S	Precision	Recall/Sensitivity	Specificity	F1 score	Accuracy
Ceftiofur	0	1	24	57	0.00	0.00	0.70	N/A	0.70
Enrofloxacin	0	15	0	67	N/A	0.00	1.00	N/A	0.82
Florfenicol	10	18	12	42	0.45	0.36	0.78	0.40	0.63
Gentamicin	34	0	41	7	0.45	1.00	0.15	0.62	0.50
Neomycin	75	2	0	5	1.00	0.97	1.00	0.99	0.98
Spectinomycin	25	1	50	6	0.33	0.96	0.11	0.50	0.38
Sulfadimethoxine	22	52	2	6	0.92	0.30	0.75	0.45	0.34
Trimeth/sulfa	6	4	26	46	0.19	0.60	0.64	0.29	0.63
Tylosin tartrate	39	27	7	9	0.85	0.59	0.56	0.70	0.59

4.2.5 Model alignment to genomes

The following steps were used as suggested by Drouin et al. to further characterize the output of Kover model analysis (Drouin, 2018). First, AR models including ER were used in a BLAST search against the family *Pasteurellaceae* (NCBI:txid712). Alignments to NCBI reference genomes were not always high quality. The best alignments (as determined through e-value, length, and percent coverage of query to subject) to the reference sequences were called reference alignments. Low e-values ($e < 10^{-50}$) and total query length at 100% coverage were considered as the reference alignment. In a case where the e-value was higher than 10^{-50} and/or coverage was less than 100%, the best alignment possible was considered. Sequence query information can be found in Appendix B.3. BLAST results identified as the reference alignment were examined to identify genes or regions of interest in the proximity of the reference alignment.

Second, the reference alignments were downloaded from NCBI as GenBank files so ER could be aligned to the reference sequence using UGENE (Okonechnikov et al., 2012; v40.1). The UGENE function "Find Pattern" was used to align the model rule sets to the matching reference alignment (i.e., model 1 for ampicillin with GenBank file from model 1 ER). Through this process a filtered alignment file was saved in fasta format for use in later steps.

4.2.6 Annotation and Target selection

OA and PA genomes were annotated using Rapid Annotation and Subsystem Technology (RAST) (Aziz et al., 2008; Brettin et al., 2015; Overbeek et al., 2013). To determine the potential function of model rules, a BLAST search was performed within RAST using the filtered alignments as the query and the OA and PA isolate genomes as the subject. For each model, if the model type was "presence", either conjunction or disjunction, the genomes considered to be in the resistant group were the subject (metadata = 1). If the model type was "absence", either conjunction or disjunction, the genomes considered to be sensitive were the subject (metadata = 0). Resulting output was searched not only to determine the sequence location within the genomes, but to examine the surrounding areas on the contigs for genes of interest.

4.2.7 Data availability

Genome sequence and coverage data, as well as AST results of OA isolates are available through their published work (Owen et al., 2017). Sequence reads from PA isolates can be found at NCBI SRA and assembled genomes are available in the NCBI Genome Repository (BioProject PRJNA824533). Bioinformatics scripts as well as input and output files for Kover and CARD are available at github.com/clwickwa/BovineRespiratoryDisease/Multispecies/scripts.

4.3 Results

4.3.1 Sequencing and assembly

The total length of PA genome assemblies, apart from A196714 (discussed below), was within the expected range of genome sizes for each species (Appendix B.1). Genome coverage was around 100 times and N50 typically ranged from 100-300 kilobase pairs (kb) with one *P. multocida* assembly (A198640) having an N50 of over 500 kb. In general, PA isolate assemblies had similar total lengths as OA isolate assemblies, indicating the genome sizes are within the expected range (Owen et al., 2017). However, the quality metrics of OA assemblies, coverage, max contig length, and N50 were much lower than those of the PA assemblies, and several OA assembles had >100 contigs. This indicates lower quality assemblies than those of the PA isolates.

PA isolate A196714 had double the expected assembly length for a *Pasteurella* spp. genome – *Pasteurella multocida* genomes should be about 2.32Mb, whereas isolate A196714 had 4.67Mb. The authors chose metagenomic taxonomic identification program Kaiju (Menzel et al., 2016) for the consideration of a co-culture. Isolate A196714 was found to contain a co-culture of *P. multocida* and *Bibersteinia trehalosi.* The isolate was still used in analysis as *B. trehalosi* is also in the family *Pasteurellaceae* and is often found as a less common opportunistic pathogen in BRD cases (Andrés-Lasheras et al., 2022). However, it should be noted that the assembly could be duplicated in places where the two species contain similar sequences.

4.3.2 Susceptibility testing and concordance using antibiotic resistance gene annotation

Antibiotic resistance genes in OA isolates found in this study were largely the same ARGs found in the Owen study (Owen et al., 2017) with a few differences. The following genes were

found in both Owen et al and the current study: *tetH, aph3-Ia, bla$_{ROB-1}$, sul2, erm42, ermF, floR,* and *dfrA14*. Additionally, the genes *aph(3``)-Ib* and *aph(6)Id* found in the current study through CARD (Alcock et al., 2019) are listed as *strA* and *strB*, respectively, in Owen et al. There is only one gene that was found in the study by Owen et al that was not found in the current study, *cat2*. While *catII* was not identified by CARD-RGI in our annotations, there was a gene found from the same class of chloramphenicol acetyltransferases, listed as *Campylobacter coli* chloramphenicol acetyltransferase (*Ccol_ACT_CHL*). Genes identified in the current study, but not in the study by Owen et al. (2017), include: two aminoglycoside resistance genes, *aadA25* and *aadB*; a second tetracycline resistance gene, *tetD*; the macrolide-lincosamide-streptogramin (MLS) resistance genes *mphE* and *msrE*; and beta-lactam resistance genes *bla$_{OXA-2}$* and *bla$_{ROB-2}$*.

The three bacterial species had different phenotypic resistance patterns, i.e., some antibiotics were found to have more resistant isolates of one species than another. MH had the most isolates with phenotypic resistance to penicillin and ampicillin and had the highest number of isolates with beta-lactam resistance genes (Table 4.2). PM had the most isolates with phenotypic resistance and ARGs for MLS antibiotics. No HS isolates were found to have a unique ARG (Appendix B.3). Phenotypic resistance to ceftiofur, a common antibiotic used for treating BRD, was found in one isolate (HS) and no specific ARGs for ceftiofur were found in any isolate (Table 4.2). Three ARGs, *aadB*, *Ccol_ACT_CHL*, and *bla$_{ROB-1,}$* were found only in MH isolates and one ARG, *bla$_{ROB-2}$*, was found in a single PM isolate. *M. haemolytica* isolates showed a diversity of phenotypic and genotypic resistance. Isolates from PA showed phenotypic resistance in at least in one isolate to all tested antibiotics with the exception of ceftiofur (Appendix B.2). Over 90% of isolates were resistant to clindamycin, neomycin, sulfadimethoxine, and tylosin, while less than 20% of isolates were resistant to ceftiofur, gentamicin, enrofloxacin, tiamulin and trimethoprim/sulfadimethoxine (TMS). More than 90% of *M. haemolytica* isolates had ARGs conferring resistance to aminoglycoside (*aph(3`)-Ia*) and tetracycline (*tetH*) antibiotics (Appendix B.4) while over 50% had the aminoglycoside resistance genes *aph(3``)-Ib* and *aph(6)-Id* as well as beta-lactam resistance gene *bla$_{ROB-1}$*. Fewer than 30% of *M. haemolytica* isolates had ARGs *erm(42)* and *sul2*, MLS and sulfonamide resistance, respectively, and less than 20% of isolates had resistance genes *floR* and *Ccol_ACT_CHL* (phenicols), *mphE* and *msrE* (MLS), *aadA(25)* and *aadB* (aminoglycoside), and *bla$_{OXA-2}$* (beta-lactam).

Among all the *P. multocida* isolates, phenotypic resistance was found in at least one isolate to all tested antibiotics except ceftiofur. However, less than 50% of isolates were resistant to nine of the 16 antibiotics (Appendix B.2). The MLS resistance gene *erm42* and tetracycline resistance gene *tetH* were found in more than 80% of isolates (Appendix B.4). Three aminoglycoside resistance genes, *aph(3``)-Ib*, *aph(6)-Id*, and *aph(3`)-Ia* were found in more than 50% of the *P. multocida* isolates (56%, 68%, and 91%, respectively). Less than 30% of the *P. multocida* isolates had resistance genes *dfrA14*, *tetD*, *ermF*, *mphE*, *msrE*, bla_{ROB-2}, *floR*, and *sul2*.

H. somni isolates displayed phenotypic resistance to very few antibiotics. Phenotypic resistance to neomycin was found in all *H. somni* isolates (Appendix B.2). While 94% of isolates were phenotypically resistant to sulfadimethoxine, corresponding TMS resistance genes were found in just under 40% of the *H. somni* isolates. Twelve of the sixteen tested antibiotics had less than 30% of isolates with resistance. Like the phenotypic resistance profiles in *H. somni*, less than 50% of *H. somni* isolates had resistance genes found in the other species. There were only three resistance genes – *erm42*, *tetH*, and *aph(3`)-Ia* – that were found in greater than 50% of the *H. somni* isolates (Appendix B.4).

Overall concordance between resistance gene and phenotypic resistance was varied depending on the antibiotic (Table 4.1). The lowest concordance was seen in sulfadimethoxine and spectinomycin with 34 and 38%, respectively. Low concordance for sulfadimethoxine was marked by a high number of false negatives (phenotypically resistant and genotypically sensitive) with 52 isolates having no sulfonamide ARG found but being phenotypically resistant to sulfadimethoxine (Table 4.1).

Table 4.2 Antibiotic resistance genes and their corresponding susceptibility testing MICs

			Breakpoint (mg/mL)	
Antibiotic Class	Resistance genes	Antibiotic Tested	Susceptible	Resistant
MLS	*erm(42), ermF, mphE, msrE*	Clindamycin	≤ 2	≥ 4
		Tilmicosin	≤ 8	≥ 16
		Tulathromycin	≤ 16	≥ 32
		Tylosin tartrate	≤ 16	≥ 32
Phenicol	*floR, Ccol_ACT_CHL*	Florfenicol	≤ 4	≥ 8
Tetracyclines	*tet(H), tet(D)*	None tested		
Aminoglycoside	*aadB, aph(6)-Id (strB), aph(3")-Ib (strA), aadA25, aadB*	Gentamicin	≤ 4	≥ 8
		Neomycin	≤ 4	≥ 8
		Spectinomycin	≤ 32	≥ 64
Trimpethoprim	*dfrA14*	Trimethoprim+ sulfamethoxazole	≤ 2	≥ 4
Sulfonamide	*sul2*	Sulfadimethoxine	≤ 256	≥ 512
Beta-lactam	*blaROB-1, blaOXA-2, blaROB-2*	Ampicillin	≤ 4	≥ 16
		Ceftiofur	≤ 2	≥ 4
		Penicillin	≤ 0.5	≥ 1
Fluoroquinolone	*none found*	Danofloxacin	≤ 0.25	≥ 0.5
		Enrofloxacin	≤ 1	≥ 2
Pleuromutilin	*none found*	Tiamulin	≤ 16	≥ 32

Conversely, low concordance for spectinomycin was the result of high false positives (phenotypically sensitive and genotypically resistant) where 50 isolates were found with aminoglycoside resistance genes but were phenotypically sensitive to spectinomycin. Ampicillin, penicillin, and neomycin had the highest concordance rates at 87%, 90%, and 98%, respectively. Clindamycin, tylosin, tilmicosin, and tulathromycin (MLS antibiotics) all had below 65% concordance. The concordance rates for clindamycin, tilmicosin, and tylosin were the result of 25-40% of the isolates being false negatives whereas the concordance rate for tulathromycin was the product of 40% of the isolates being false positives.

4.3.3 Model statistics, alignment, and annotation

Multiple MIC resistance breakpoints were considered for each antibiotic and the resulting resistant/sensitive metadata was used as input to Kover (Table 4.3). The results from Kover confirmed that the models at the previously established breakpoint MIC had the best overall model, indicating Kover behaved the way we expected and generated reasonable models.

Models included: 1) whether there were multiple rules associated with AR, 2) if the model was a combination of all rules (conjunction) or individual rules (disjunction), as well as 3) if it was the presence (rule found in AR isolates) or absence (rule found in AS isolates) of the rule that was considered as the AR determinant. Models at breakpoint MICs resulted in both presence and absence, as well as both conjunction and disjunction models (Table 4.4). Model error rates and statistics can be found in Table 4.3 for the seven antibiotics tested. Out of the seven antibiotics, machine learning improved the concordance for five antibiotics (Table 4.4). However, it is important to note that some of the models are under fit, meaning the training error rate is much lower than testing error rate (Table 4.3).

As all isolates with an MIC interpretation of ‘intermediate’ were considered resistant, no minor errors – defined as the misclassification of isolates within one interpretive category (e.g., susc-int, int-resist) – were recorded. The results, therefore, are given as overall error (OE), defined as the total of all misclassifications; major errors (ME), defined as true susceptible isolates classified as resistant; and very major errors (VME), defined as true resistant isolates classified as susceptible. ME ranged from 0-22% of isolates while VME ranged from and 0-25% of isolates (Table 4.5). Misclassification of isolates by a model was seen for every antibiotic except penicillin, for which the OE was 0% (Table 4.5). Tiamulin had no ME, though the model did have the highest

rate of VME at 25%. Conversely, tulathromycin and tilmicosin had no VME, however tilmicosin had the second highest rate of ME (14%; Table 5). The highest ME was seen from the danofloxacin model with 22% of the isolates misclassified as resistant.

Taxonomic classification of reference genomes found from NCBI BLAST varied between antibiotics as well as rulesets (Table 4.4). However, the search often returned alignments to a single species for each model (Appendix B.5). All rules were constrained to a BLAST search against the family *Pasteurellaceae*, as searching more broadly resulted in no or low confidence BLAST hits (e-values close to 1). Results from individual antibiotic models are discussed over the next several paragraphs.

Table 4.3 Kover error data of BRD antibiotics. Boxed summary information represents the currently established MIC breakpoint value (CLSI reference) for each antibiotic

		Training Data				Testing Data			
ABX	MIC	Susc	Res	Error	F1-score	Susc	Res	Error	F1-score
DANO	0.25mg	31	24	0.091	0.88	14	13	0.22	0.77
	0.5mg	36	19	0.11	0.84	19	8	0.33	0.53
TIL	8mg	14	41	0.091	0.94	6	21	0.26	0.84
	16mg	20	35	0.091	0.93	8	19	0.15	0.9
	32mg	36	19	0.73	0.9	16	11	0.11	0.87
	64mg	36	19	0	1	18	9	0.3	0.67
	> 64mg	45	10	0.11	0.57	24	3	0.26	N/A
TULA	8mg	14	41	0.14	0.9	5	22	0.3	0.83
	16mg	36	19	0.018	0.97	17	10	0.48	0.43
	32mg	45	10	0.055	0.87	25	2	0.074	0.67
	≥ 64mg*								
AMP	0.5mg	35	20	0.054	0.93	20	7	0.48	0.43
	1mg	40	15	0.073	0.85	21	6	0.19	0.54
	4mg	41	14	0.073	0.83	21	6	0.074	0.86
	16mg	42	13	0.036	0.92	21	6	0.11	0.73
	32mg	45	10	0.018	0.95	21	6	0.15	0.75
CLIN	2mg	9	46	0.055	0.98	4	23	0.037	0.95
	4mg	13	42	0.02	0.99	5	22	0.11	0.93
	8mg	14	41	0.036	0.98	5	22	0.037	0.98
	16mg	24	31	0.036	0.97	13	14	0.18	0.83
	>16mg	36	19	0.018	0.97	16	11	0.11	0.87
PEN	0.25mg	16	39	0.11	0.93	13	14	0.26	0.8
	0.5mg	36	19	0.036	0.95	20	7	0.18	0.62
	1mg	38	17	0.11	0.79	20	7	0	1
	4mg	40	15	0.055	0.89	21	6	0.15	0.67
	8mg	41	14	0.036	0.93	21	6	0.22	0.67
	16mg	44	11	0.036	0.92	21	6	0.18	0.71
TIA	4mg	13	42	0.036	0.98	5	22	0.11	0.94
	8mg	14	41	0	1	6	21	0.074	0.95
	16mg	21	34	0.018	0.98	11	16	0.15	0.87
	32mg	38	17	0.018	0.97	19	8	0.26	0.22

* Not enough isolates in group ≥ 64 mg (established breakpoint) to determine model

Table 4.4 KOVER model for breakpoint MIC. Model type: presence (P) or absence (A) of the ruleset predicts resistance; Importance: weighted value for rule – designates how often the rule was found in the model; For multiple rulesets - Conjunction: all rulesets pre predict resistance (logical AND) or Disjunction: one ruleset needs to be present or absent to predict resistance (logical OR).

ABX	MIC	Model type (Importance)		# equiv rules	Annotation
DANO	0.5mg	Conjunction	P (0.83)	25	IS30-like transposase
			A (0.22)	31	*murA* – peptidoglycan synthesis
TIL	16mg		A (1.00)	128	23S rRNA (multiple regions)
TULA	≥ 32mg	Conjunction	P (0.71)	17	Non-coding region between *tetH* and *tetR*
			A (0.40)	2	23S rRNA (*Pasteurella* spp.)
AMP	16mg	Conjunction	P (0.98)	73	histidinol transaminase
			P (0.93)	66	mobile element protein
CLIN	4mg	Disjunction	A (0.95)	87	23S rRNA
			P (0.93)	2	proline tRNA
PEN	1mg		P (1.00)	31	histidinol dehydrogenase
TIA	32mg		P (1.00)	3	CRISPR array

The predictive model for phenotypic danofloxacin resistance had two rules and was classified as a conjunction model (Table 4.4). The presence of the first rule (25 ER) was annotated as mobile element protein that matched IS30-like transposase in BLAST. Rule two (31 ER) was typed as absence and was annotated as *murA*, a key enzyme for the synthesis of peptidoglycan (Table 4.4). A total of 15 isolates were misclassified by this model – six from the training data set and nine from testing. Of the 15 misclassified isolates, eight were *MH*, six were *PM*, and one was *HS.* Three of the *MH* isolates were misclassified as danofloxacin susceptible (one training, two testing) and five were misclassified as being danofloxacin resistant (one training, four testing). There were four *PM* isolates misclassified as danofloxacin susceptible (two training, two testing) and two with resistant phenotypes (one training, one testing). The single *HS* isolate was considered danofloxacin resistant by the model (training). Additionally, half of the isolates with an MIC of 1 mg/mL (resistant) were misclassified.

The model for tulathromycin resistance had two conjunction rules (Table 4.4). The presence of the first rule (17 ER) matched an area between *tetH* and *tetR*, tetracycline resistance genes. The absence of rule two (2 ER) was annotated as a region of 23S rRNA gene specific to *Pasteurella* spp. (Appendix B.5). Five isolates were misclassified by the model for tulathromycin resistance (three training, two testing). One *MH* isolate was misclassified as resistant from

(training). No *PM* isolates were misclassified. Four misclassifications were from *HS* isolates, all with tulathromycin susceptible phenotypes misclassified as resistant (two training, two testing). Isolates from all species were predicted to be both resistant and sensitive.

Tilmicosin resistance prediction resulted in a single rule (128 ER); the absence of multiple variable regions of the 23S rRNA gene, a gene important as the target of MLS antibiotics, the class to which tilmicosin belongs (Table 4.4). On the other hand, 23S rRNA is important in differentiating species, including those of the family *Pasteurellaceae*. Of the nine isolates misclassified by the predictive model for tilmicosin resistance, five were from training and four from testing datasets. All five of the *MH* isolates were misclassified to be resistant to tilmicosin. Interestingly, no *MH* isolates with phenotypic susceptibility to tilmicosin were correctly identified during training or predicted during testing. Two *PM* isolates were misclassified as resistant (testing). One phenotypically susceptible and resistant *HS* isolate each were misclassified (training).

The model for ampicillin resistance had two rules and was classified as a conjunction model. The presence of rule one (73 ER) was annotated as a histidinol transaminase, an important enzyme for the synthesis of several amino acids (e.g., histidine, tyrosine, and tryptophan) while the presence of rule two (66 ER) was annotated as a mobile element protein (table annotation info). There were a total of five misclassified isolates by the ampicillin model, two in the training data set and three in the testing data set. Three of the *MH* isolates (one training, two testing) were misclassified as susceptible by the model while one *MH* isolate (testing) was misclassified as resistant. The single phenotypically resistant *H. somni* isolate (training) and one of the two *P. multocida* isolates (training) with a resistant phenotype was correctly identified by the model, whereas the remaining *PM* isolate (training) with ampicillin resistant phenotype was misclassified as susceptible. No *HS* or *PM* isolates with antibiotic susceptible phenotypes were misclassified (Table 4.6) and the model effectively identifies susceptible and resistant *MH* isolates.

Table 4.5 Comparison of error rates and F1-scores between resistance database query (ARG) and machine learning (ML) at the established breakpoint for each tested antibiotic

	ARG		ML	
Antibiotic	Error Rate	F1-score	Error Rate	F1-score
Danofloxacin	33%	N/A	33%	0.53
Tilmicosin	39%	0.68	15%	0.9
Tulathromycin	41%	0.41	7.4%	0.67
Ampicillin	13%	0.76	11%	0.73
Clindamycin	49%	0.68	11%	0.93
Penicillin	10%	0.84	0%	1
Tiamulin	30%	N/A	26%	0.22

The model for penicillin resistance had a single rule (31 ER), the presence of which was annotated as histidinol dehydrogenase, a different enzyme in the same amino acid synthesis pathway as ampicillin (Table 4.4). There were six total misclassifications from the penicillin model, all in the training data set. Three *MH* and *PM* isolates each were misclassified as susceptible. Additionally, the three *PM* isolates were the only *PM* isolates with phenotypic penicillin resistance. No *HS* isolates were misclassified; however, they were all phenotypically susceptible. It appears the model effectively identifies susceptible and resistant *MH* isolates, while assuming all PM and HS isolates are susceptible.

The tiamulin model had a single rule, for which the presence (3 ER; table 4.4) matched a CRISPR array with each ER being repeated across a 5-7 kilobase region (Appendix B.2). Eight isolates were misclassified by the tiamulin resistance model: one error from the training dataset and seven from testing. Three *MH* isolates were misclassified as susceptible, which were the only *MH* isolates with phenotypic resistance to tiamulin. The single training error was a *PM* isolate, and the remaining four testing errors were *PM* isolates misclassified as susceptible. No *HS* isolates were misclassified; however, they were all phenotypically susceptible.

The clindamycin model had two rules and was classified as a disjunction model. The absence of rule one (87 ER) matched 23S rRNA in *Pasteurellaceae*, but the rule was split into two sections of the 23S rRNA gene which was surrounded by other rRNA (Table 4.4). The presence of rule two (2 ER) matched a proline tRNA with the specific anticodon TGG. A total of seven isolates were misclassified by the clindamycin model, four in the training data set and three in

testing data set. The training errors were an *MH* isolate and two HS isolates, misclassified as resistant, and one PM isolate, misclassified as susceptible. Two of the three testing errors were *PM* isolates misclassified as susceptible. The remaining testing error was an *HS* isolate misclassified as resistant.

4.3.4 Alignment similarities and overall error

The sets of equivalent rules for each model were used to create alignments, however not every isolate had every equivalent rule. Pencillin was the only antibiotic for which the percent of similarity between the filtered alignment (query) and isolate genome (subject) was below 80%. Penicillin, with only one model, had the lowest OE of all the antibiotics at 0% (Table 4.3) but some isolates had 55% percent similarity with the filtered alignment (Appendix B.2). Interestingly, though all *HS* isolates were phenotypically susceptible to penicillin and the only *PM* isolates with phenotypic resistance to penicillin were misclassified as susceptible by the predictive model, the penicillin model had no testing errors (Table 4.6), the highest accuracy, and highest F1-score (table 4.3).

The rule for tiamulin was a CRISPR array – a section of repeated bases with "spacers" of DNA sequences from previously encountered foreign genetic elements (Garrett, 2021). As the spacers are not always the same for every species, or even every isolate of the same species, determination of similarity to the filtered alignment was not considered necessary for this study. However, as CRISPR/Cas systems can inhibit horizontal gene transfer (Wheatley & MacLean, 2021), it would be interesting to examine the types of sequences encountered and captured within the cassettes.

While the OE of antibiotic resistance prediction for the majority of antibiotics was improved using machine learning (Table 4.5), it is important to examine the individual rules and their annotations if present. Four of the seven antibiotics examined in this study were associated with models that had multiple rules. Of these, two had at least one rule with reference alignments to only one species – danofloxacin rule 2 and tulathromycin rule 2 only had reference alignments to *P. multocida*. Additionally, the single rule for the tiamulin model only had reference alignments to *P. multocida*. In each of these three cases, the resulting annotation was a gene (danofloxacin and tulathromycin) or non-coding region (tiamulin) that was specific to *P. multocida*.

Table 4.6 Kover models with classification of each species

		Training						Testing					
		Correct			Error			Correct			Error		
		MH	PM	HS	MH	PM	HS	MH	PM	HS	MH	PM	HS
Danofloxacin	Res	6	9	1	1	1	1	1	4	0	2	1	0
	Susc	12	11	10	1	2	0	5	2	6	4	2	0
Tilmicosin	Res	20	13	1	0	0	1	7	12	0	0	0	0
	Susc	0	4	12	3	0	1	0	1	3	2	2	0
Tulathromycin	Res	7	2	1	0	0	0	2	0	0	0	0	0
	Susc	15	17	10	1	0	2	7	13	3	0	0	2
Ampicillin	Res	10	1	0	1	1	0	4	0	0	2	0	0
	Susc	9	20	13	0	0	0	5	10	5	1	0	0
Clindamycin	Res	19	21	2	0	1	0	12	8	0	0	2	0
	Susc	0	1	11	1	0	2	0	0	4	0	0	1
Penicillin	Res	11	0	0	3	3	0	7	0	0	0	0	0
	Susc	8	16	14	0	0	0	3	13	4	0	0	0
Tiamulin	Res	0	17	0	0	1	0	0	0	0	3	4	0
	Susc	19	0	18	0	0	0	10	10	0	0	0	0

However, it was not always the case that antibiotics with singularly skewed species resistance resulted in a seemingly biased model. For example, the model for clindamycin, with a 38% improvement in error rate compared to ARG comparison alone (Table 4.5), was built from data where 16 of 18 *HS* isolates and only one isolate of MH and PM, each, were found in the phenotypically susceptible group (Table 4.6). The clindamycin rules matched sections of the 23S rRNA gene, but the regions matched were conserved across many genera in the family *Pasteurellaceae* (Table 4.4). Similarly, the model for penicillin resistance matched an amino acid synthesis gene and was built from data where 84% of the phenotypically resistant isolates were MH, the remaining 16% were PM, while all HS isolates were phenotypically susceptible to penicillin (Table 4.4).

4.4 Discussion

In the current study, machine learning was used to determine sequences predictive of antibiotic resistance from bacterial genome assemblies and phenotypic AST data. Presently, two methods of detecting antibiotic resistance are commonly used in veterinary laboratories: 1)

antibiotic susceptibility testing (AST) and molecular assays with bacteria isolated from sick or dead animals, or 2) molecular assays to determine presence of antibiotic resistance genes (ARG) within unpurified samples using polymerase chain reaction (PCR) amplification of target ARGs. In the case of BRD, however, these two methods have distinct issues that make their use impractical for accurate identification of resistance. First, bacterial culture and AST can be slow, requiring days for fast-growing species such as *E. coli* or *Salmonella* spp. or weeks for slow-growing species such as *Mycobacterium tuberculosis* (Beste et al., 2009; Neidhardt, 1996). Second, ARGs must be discovered and catalogued, thus determination of antibiotic resistance genes is dependent on previous determination of gene function. High genotype-phenotype concordance has been observed in well-studied organisms which also have assembled genomes from many isolates of the species. For instance, a study on the concordance of ARGs and phenotypic AR of non-typhoidal *Salmonella enterica* found 97.8% of approximately 3,400 isolates had AR genotype and phenotype that agreed (Neuert et al., 2018). In contrast, when studying 64 isolates from three pathogens associated with BRD, Owen et al (2017) found less than 75% concordance between AR genotype and phenotype. In the case of BRD pathogens, there are additional genomes sequences and assemblies that are publicly available; however, these genomes lack accompanying phenotypic antibiotic resistance information which limits ability to determine concordance with a larger sample size.

4.4.1 Antibiotic resistance patterns across BRD pathogens

Initial examination of the phenotypic resistance profile revealed HS isolates to be susceptible to most antibiotics (Table 4.2). In contrast, ceftiofur was the only antibiotic for which there were no resistant isolates of PM and MH. The profile for resistance in PM and MH showed the isolates are highly resistant to sulfadimethoxine, the aminoglycoside antibiotics clindamycin and neomycin, and the MLS antibiotics tilmicosin and tylosin. Of those antibiotics, tilmicosin is the only antibiotic listed for which there is active use of the antibiotic to treat BRD under the trade name Micotil. As the majority of PM (25/32 isolates) and MH (27/32 isolates) were phenotypically resistant to tilmicosin, it may be necessary to avoid use of the antibiotic as a primary treatment if the causative agent is not known.

Curiously tulathromycin, an antibiotic also actively used in the treatment of BRD (trade name Draxxin), had very low numbers of isolates with phenotypic resistance in all three pathogens.

Previous studies on the use of tulathromycin in prevention and treatment of BRD have shown tulathromycin resistance to be present in many MH isolates prior to use. As we do not have antibiotic use information on all the animals from which the isolates in the current study were sourced, it is unknown whether the animals were treated with tulathromycin.

4.4.2 Models with biologically significant AR predictors

The majority of models contained, or were entirely comprised of, rules that had annotations matching genes or other coding sequences. Ampicllin model rule 1 and the penicillin model both matched genes in the histidine/purine biosynthesis pathway (Table 4.4). Ampicillin model rule 1 matched histidinol phosphate transferase (*hisC*; EC:2.6.1.9.) which catalyzes the transfer of phosphate between L-histidinol-P and imidazole-acetol-P. The penicillin model ruleset matched histidinol deydrogenase (*hisD*; EC:1.1.1.23) which catalyzes the removal of a hydrogen from L-histidinol, resulting in L-histidinal, an immediate precursor to L-histidine. The link between histidine biosynthesis and beta-lactam resistance could be the role of histidine metabolism in cell wall synthesis. A well studied amino acid, histidine is a proteogenic amino acid that is also utilized in the metabolism of other amino acids and amino sugars, including those used in the formation of peptidoglycan (Bender, 2012; Zeng & Lin, 2013). Additionally, the his operon has been implicated in increased multidrug resistance following duplication of the operon (Brenner & Ames, 1971; Dunn et al., 2021).

Some model rules were annotated as genes that could be associated with the potential targets of the antibiotics – danofloxacin model rule 2 and tulathromycin model rule 2, as well as ampicillin model rule 1 and the penicillin model discussed above (Table 4.4). Interestingly, the second danofloxacin rule matched a cell wall synthesis gene and as a fluoroquinolone, danofloxacin targets cell replication machinery (DNA gyrase and topoisomerase IV). The second rule for tulathromycin, a macrolide antibiotic, matched one of the targets of macrolides, the 23S rRNA subunit. Since most of these rules seem to be related to the target of the antibiotic, resistance might be conferred through antibiotic target alteration, potentially preventing the inhibition of the bacteria. Further characterization is required to confirm this hypothesis.

In the current study, there were two models with rules annotated as MGEs. Danofloxacin model rule 1 and ampicillin model rule 2 had matching sequences to mobile element proteins that were annotated as insertion sequences (IS) of an integrative-conjugative element (ICE) in the

reference genomes (Table 4.4). In both models the IS was annotated as "IS30-like element ISApl1 family transposase," an IS first found in another member of the *Pasteurellaceae* family*, Actinobacillus pleuropneumoniae.* In the danofloxacin model, there were many types of antibiotic resistance genes on the same contig as the IS including aminoglycoside, sulfonamide, and phenicol as well as potential conjugation machinery, suggesting the importance of resistance genes to the survival of these pathogens. Interestingly, in the ampicillin model most of the surrounding genes were other MGE sequences including integrases and "mobile element" or "mobilization" proteins.

As MGEs can be associated with many different genes, the confidence of targeting a specific ARG or sequence is difficult. The ICEs found in the current study have been identified as having the same conserved backbone in MH, PM, and HS (Andrés-Lasheras et al., 2022; Eidam et al., 2015; Lubbers & Hanzlicek, 2013; Michael et al., 2012; Owen et al., 2017). As additional evidence as to why targeting an MGE would be challenging, the regions surrounding the MGE sequences were conjugation machinery (type IV pilus system) and other mobile genetic elements (transposases and integrases) that are often the conserved sequences in MGEs, as well as antimicrobial resistance genes (e.g., copper resistance and phenicol resistance) that are not always consistent between isolates (Beker et al., 2018; Frost et al., 2005; Partridge et al., 2018).

To use the MGEs as predictor, there needs to be further research on the types of AR co-occurring with the MGEs in the three species from the current study. There were several models for which model rules had annotations other than potential antibiotic targets (e.g., 23S rRNA) or ARG (e.g., *tetR*). These models could be targeting a proxy gene – not a gene conferring the antibiotic resistance in question but one found to be frequently associated, or co-occur, with resistant genes (Johnson et al., 2016). Genes and sequences surrounding the model rules of danofloxacin, tulathromycin, and ampicillin (Table 4.4) were other AMR genes, e.g., metal resistance genes, and MGEs. Co-occurrence of AMR genes is not a new concept, and further research into what types of genes co-occur in the BRD pathogens may improve confidence that the non-ARG sequences could be used as proxies for co-occurring ARGs.

4.4.3 Considerations for the use of non-coding sequences as AR predictors

Tulathromycin model rule 1 matched a non-coding sequence and it is currently unclear how non-protein-coding sequences would confer an antibiotic resistance phenotype. We have three hypotheses: the model rule sequences are (i) merely proxy sequences for resistant bacterial clades,

(ii) epigenetic controls, such as methylation and acetylation, resulting in a "dormant" survival phenotype (Ghosh et al., 2020; Riber & Hansen, 2021), or (iii) regulatory sequences or binding sites (e.g., sigma factors) facilitating stress-response mechanisms during exposure to antibiotics (Woods & McBride, 2017; Yoo et al., 2016).

4.4.4 Future research directions

The functional annotations of the rule sequences for each model in the current study revealed a wide variety of coding and non-coding sequences identified as the predictive sequences of phenotypic resistance. Ideally, the rules would have matched coding sequences that could be identified as a novel resistance gene. Novel resistance genes often have homologs with the same functions or might be annotated as hypothetical proteins, but proper annotation of homologous genes may not occur (Bengtsson-Palme, 2018). Additional testing (bioinformatic, biochemical, or multi-omic) is required to confirm if the rules produce functional genes and the importance of the non-coding sequences.

Though all three of the isolate genera belonged to the family *Pasteurellaceae*, there are still enough differences in their genomes to consider them separate genera and species. An example of phenotypic differences is in the phenotypic resistance profile of HS isolates compared to MH and PM (Appendix B.2); HS isolates typically have fewer antibiotics to which they are phenotypically resistant. Additionally, the pan-genome of the three taxa used in the current study may have included too many species-specific sequences to determine AR markers across all three species. A recent methodology using the pan-genome in antimicrobial resistance prediction improved the likelihood of correct predictions of four species with open pan-genomes (Yang & Wu, 2022), suggesting the ability to mine the pan-genomes of the species in the current study for enhanced prediction.

Due to the potential limitation of using the pan-genome of these three species, it is imperative that further analysis of individual BRD pathogens be carried out. However, the pathogens in the current study have very few available isolates in the NCBI database that have phenotypic resistance data along with sequenced and/or assembled genomes. For *Mannheimia haemolytica,* there are 2,023 isolates with sequence reads and 207 isolates with assembled genomes available, while *Pasteurella multocida* has 1,244 isolates with sequence reads and 310 with assembled genomes available. However, the majority of the *M. haemolytica* and *P. multocida*

isolates do not have AST data publicly available. For example, there are only 18 available isolates of *Histophilus somni* from NCBI with AST and genomic data. In addition to the low numbers of paired sequence and antibiotic resistance phenotype data, not all the available isolates come from a relevant environment, i.e., ruminant animals, with emphasis on beef or dairy production.

4.5 Conclusions

Using machine learning, the rate of predicting whether an isolate was resistant to a particular antibiotic was increased. Prediction of phenotypic resistance to the BRD antibiotics tilmicosin and tulathromycin improved 24 and 33%, respectively, using machine learning and that pattern was seen for all antibiotics except danofloxacin. However, the functional annotations of the model rule sequences were not clear indicators of the mechanism of resistance. There were antibiotics for which the majority of sensitive or resistant isolates were from only one of the three species (e.g., tiamulin), creating likely bias in the sequences found. Additionally, there were AR determinants found for mismatching antibiotic classes (e.g., danofloxacin matching the *murA* cell wall synthesis gene). The potential species indicators and mismatching resistances suggests there may be more to optimize in identifying resistance determinants in these pathogens.

The lack of available paired genomic and phenotypic data is a current limitation for the determination of AR using the genome in BRD pathogens. Though the machine learning method used in this study increased the accuracy of AR prediction, there is still room for improvement. It is necessary to add to the current repository of sequenced BRD pathogens with AST data to further assess the variability in BRD pathogen genomes. This will allow more comparisons of isolates across different geographical areas and times, as well as increase the knowledge of the core- and pan-genomes of BRD pathogens.

4.6 References

Adam, M., Murali, B., Glenn, N. O., & Potter, S. S. (2008). Epigenetic inheritance based evolution of antibiotic resistance in bacteria. BMC Evolutionary Biology, 8(1), 52. https://doi.org/10.1186/1471-2148-8-52

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

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., … McArthur, A. G. (2019). CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Research, gkz935. https://doi.org/10.1093/nar/gkz935

Andrés-Lasheras, S., Jelinski, M., Zaheer, R., & McAllister, T. A. (2022). Bovine Respiratory Disease: Conventional to culture-independent approaches to studying antimicrobial resistance in north america. Antibiotics, 11(4), 487. https://doi.org/10.3390/antibiotics11040487

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008). The RAST server: rapid annotations using subsystems technology. BMC Genomics, 9(1), 75. https://doi.org/10.1186/1471-2164-9-75

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology, 19(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Beker, M., Rose, S., Lykkebo, C. A., & Douthwaite, S. (2018). Integrative and Conjugative Elements (ICEs) in Pasteurellaceae species and their detection by multiplex PCR. Frontiers in Microbiology, 9, 1329. https://doi.org/10.3389/fmicb.2018.01329

Bender, R. A. (2012). Regulation of the histidine utilization (hut) system in bacteria. Microbiology and Molecular Biology Reviews, 76(3), 565–584. https://doi.org/10.1128/MMBR.00014-12

Bengtsson-Palme, J. (2018). The diversity of uncharacterized antibiotic resistance genes can be predicted from known gene variants—But not always. Microbiome, 6(1), 125. https://doi.org/10.1186/s40168-018-0508-2

Berglund, F., Österlund, T., Boulund, F., Marathe, N. P., Larsson, D. G. J., & Kristiansson, E. (2019). Identification and reconstruction of novel antibiotic resistance genes from metagenomes. Microbiome, 7(1), 52. https://doi.org/10.1186/s40168-019-0670-1

Beste, D. J. V., Espasa, M., Bonde, B., Kierzek, A. M., Stewart, G. R., & McFadden, J. (2009). The genetic requirements for fast and slow growth in Mycobacteria. PLoS ONE, 4(4), e5349. https://doi.org/10.1371/journal.pone.0005349

Brenner, M., & Ames, B. N. (1971). The histidine operon and its regulation. In Metabolic Regulation (pp. 349–387). Elsevier. https://doi.org/10.1016/B978-0-12-299255-1.50018-3

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Scientific Reports, 5(1), 8365. https://doi.org/10.1038/srep08365

Davidson, J. L., Wang, J., Maruthamuthu, M. K., Dextre, A., Pascual-Garrigos, A., Mohan, S., Putikam, S. V. S., Osman, F. O. I., McChesney, D., Seville, J., & Verma, M. S. (2021). A paper-based colorimetric molecular test for SARS-CoV-2 in saliva. Biosensors and Bioelectronics: X, 9, 100076. https://doi.org/10.1016/j.biosx.2021.100076

Drouin, A. (2018, August). Interpreting Kover Models. Kover Documentation. https://aldro61.github.io/kover/doc_interp.html

Drouin, A., Giguère, S., Déraspe, M., Marchand, M., Tyers, M., Loo, V. G., Bourgault, A.-M., Laviolette, F., & Corbeil, J. (2016). Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons. BMC Genomics, 17(1), 754. https://doi.org/10.1186/s12864-016-2889-6

Drouin, A., Letarte, G., Raymond, F., Marchand, M., Corbeil, J., & Laviolette, F. (2019). Interpretable genotype-to-phenotype classifiers with performance guarantees. Scientific Reports, 9(1), 4071. https://doi.org/10.1038/s41598-019-40561-2

Dunn, S., Carrilero, L., Brockhurst, M., & McNally, A. (2021). Limited and Strain-Specific Transcriptional and growth responses to acquisition of a multidrug resistance plasmid in genetically diverse Escherichia coli lineages. MSystems, 6(2), e00083-21. https://doi.org/10.1128/mSystems.00083-21

Eidam, C., Poehlein, A., Leimbach, A., Michael, G. B., Kadlec, K., Liesegang, H., Daniel, R., Sweeney, M. T., Murray, R. W., Watts, J. L., & Schwarz, S. (2015). Analysis and comparative genomics of ICEMh1, a novel integrative and conjugative element (ICE) of Mannheimia haemolytica. 5. https://doi.org/doi:10.1093/jac/dku361

Frost, L. S., Leplae, R., Summers, A. O., & Toussaint, A. (2005). Mobile genetic elements: The agents of open source evolution. Nature Reviews Microbiology, 3(9), 722–732. https://doi.org/10.1038/nrmicro1235

Garrett, S. C. (2021). pruning and tending immune memories: spacer dynamics in the CRISPR array. Frontiers in Microbiology, 12, 664299. https://doi.org/10.3389/fmicb.2021.664299

Ghosh, D., Veeraraghavan, B., Elangovan, R., & Vivekanandan, P. (2020). Antibiotic resistance and epigenetics: more to it than meets the eye. Antimicrobial Agents and Chemotherapy, 64(2), e02225-19. https://doi.org/10.1128/AAC.02225-19

Gremme, G., Steinbiss, S., & Kurtz, S. (2013). GenomeTools: A comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 10(3), 645–656. https://doi.org/10.1109/TCBB.2013.68

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. Bioinformatics, 29(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Johnson, T. A., Stedtfeld, R. D., Wang, Q., Cole, J. R., Hashsham, S. A., Looft, T., Zhu, Y.-G., & Tiedje, J. M. (2016). Clusters of antibiotic resistance genes enriched together stay together in swine agriculture. MBio, 7(2), e02214-15–11. https://doi.org/10.1128/mBio.02214-15

Knopp, M., Babina, A. M., Gudmundsdóttir, J. S., Douglass, M. V., Trent, M. S., & Andersson, D. I. (2021). A novel type of colistin resistance genes selected from random sequence space. PLOS Genetics, 17(1), e1009227. https://doi.org/10.1371/journal.pgen.1009227

Köser, C. U., Ellington, M. J., Cartwright, E. J. P., Gillespie, S. H., Brown, N. M., Farrington, M., Holden, M. T. G., Dougan, G., Bentley, S. D., Parkhill, J., & Peacock, S. J. (2012). Routine Use of microbial whole genome sequencing in diagnostic and public health microbiology. PLoS Pathogens, 8(8), e1002824. https://doi.org/10.1371/journal.ppat.1002824

Kurtz, S., Narechania, A., Stein, J. C., & Ware, D. (2008). A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes. BMC Genomics, 9(1), 517. https://doi.org/10.1186/1471-2164-9-517

Lubbers, B. V., & Hanzlicek, G. A. (2013). Antimicrobial multidrug resistance and coresistance patterns of Mannheimia haemolytica isolated from bovine respiratory disease cases—A three-year (2009–2011) retrospective analysis. Journal of Veterinary Diagnostic Investigation, 25(3), 413–417. https://doi.org/10.1177/1040638713485227

Marchand, M., & Shawe-Taylor, J. (2002). The set covering machine. Journal of Machine Learning Research, 3, 723–746.

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature Communications, 7(1), 11257. https://doi.org/10.1038/ncomms11257

Michael, G. B., Kadlec, K., Sweeney, M. T., Brzuszkiewicz, E., Liesegang, H., Daniel, R., Murray, R. W., Watts, J. L., & Schwarz, S. (2012). ICEPmu1, an integrative conjugative element (ICE) of Pasteurella multocida: Structure and transfer. Journal of Antimicrobial Chemotherapy, 67(1), 91–100. https://doi.org/10.1093/jac/dkr411

Mikheenko, A., Saveliev, V., & Gurevich, A. (2016). MetaQUAST: Evaluation of metagenome assemblies. Bioinformatics, 32(7), 1088–1090. https://doi.org/10.1093/bioinformatics/btv697

Murray, G. M., O'Neill, R. G., More, S. J., McElroy, M. C., Earley, B., & Cassidy, J. P. (2016). Evolving views on bovine respiratory disease: An appraisal of selected key pathogens – Part 1. The Veterinary Journal, 217, 95–102. https://doi.org/10.1016/j.tvjl.2016.09.012

Neidhardt, F. C. (1996). Escherichia coli and Salmonella: Cellular and molecular biology.

Neuert, S., Nair, S., Day, M. R., Doumith, M., Ashton, P. M., Mellor, K. C., Jenkins, C., Hopkins, K. L., Woodford, N., de Pinna, E., Godbole, G., & Dallman, T. J. (2018). Prediction of phenotypic antimicrobial resistance profiles from whole genome sequences of non-typhoidal Salmonella enterica. Frontiers in Microbiology, 9, 592. https://doi.org/10.3389/fmicb.2018.00592

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S. R., Woyke, T., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2013). Assembling genomes and mini-metagenomes from highly chimeric reads. In M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), Research in Computational Molecular Biology (pp. 158–170). Springer Berlin Heidelberg.

Okonechnikov, K., Golosova, O., & Fursov, M. (2012). Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics, 28(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2013). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Research, 42(D1), D206–D214. https://doi.org/10.1093/nar/gkt1226

Owen, J. R., Noyes, N., Young, A. E., Prince, D. J., Blanchard, P. C., Lehenbauer, T. W., Aly, S. S., Davis, J. H., O'Rourke, S. M., Abdo, Z., Belk, K., Miller, M. R., Morley, P., & Van Eenennaam, A. L. (2017). Whole-genome sequencing and concordance between antimicrobial susceptibility genotypes and phenotypes of bacterial isolates associated with bovine respiratory disease. G3;#58; Genes|Genomes|Genetics, 7(9), 3059–3071. https://doi.org/10.1534/g3.117.1137

Partridge, S. R., Kwong, S. M., Firth, N., & Jensen, S. O. (2018). Mobile genetic elements associated with antimicrobial resistance. Clinical Microbiology Reviews, 31(4), e00088-17, /cmr/31/4/e00088-17.atom. https://doi.org/10.1128/CMR.00088-17

Pascual-Garrigos, A., Maruthamuthu, M. K., Ault, A., Davidson, J. L., Rudakov, G., Pillai, D., Koziol, J., Schoonmaker, J. P., Johnson, T., & Verma, M. S. (2021). On-farm colorimetric detection of Pasteurella multocida, Mannheimia haemolytica, and Histophilus somni in crude bovine nasal samples. Veterinary Research, 52(1), 126. https://doi.org/10.1186/s13567-021-00997-9

Peel, D. S. (2020). The effect of market forces on bovine respiratory disease. Veterinary Clinics of North America: Food Animal Practice, 36(2), 497–508. https://doi.org/10.1016/j.cvfa.2020.03.008

Pehrsson, E. C., Forsberg, K. J., Gibson, M. K., Ahmadi, S., & Dantas, G. (2013). Novel resistance functions uncovered using functional metagenomic investigations of resistance reservoirs. Frontiers in Microbiology, 4. https://doi.org/10.3389/fmicb.2013.00145

Petruzzi, B., Dickerman, A., Lahmers, K., Scarratt, W. K., & Inzana, T. J. (2020). Polymicrobial biofilm interaction between Histophilus somni and Pasteurella multocida. Frontiers in Microbiology, 11, 1561. https://doi.org/10.3389/fmicb.2020.01561

Riber, L., & Hansen, L. H. (2021). Epigenetic memories: the hidden drivers of bacterial persistence? Trends in Microbiology, 29(3), 190–194. https://doi.org/10.1016/j.tim.2020.12.005

Sandal, I., & Inzana, T. J. (2010). A genomic window into the virulence of Histophilus somni. Trends in Microbiology, 18(2), 90–99. https://doi.org/10.1016/j.tim.2009.11.006

Theurer, M. E., Johnson, M. D., Fox, T., McCarty, T. M., McCollum, R. M., Jones, T. M., & Alkire, D. O. (2021). Bovine respiratory disease during the mid-portion of the feeding period: Observations of frequency, timing, and population from the field. Applied Animal Science, 37(1), 52–58. https://doi.org/10.15232/aas.2020-02089

Wang, Y., & Guan, L. L. (2022). Translational multi-omics microbiome research for strategies to improve cattle production and health. Emerging Topics in Life Sciences, 6(2), 201–213. https://doi.org/10.1042/ETLS20210257

Wheatley, R. M., & MacLean, R. C. (2021). CRISPR-Cas systems restrict horizontal gene transfer in Pseudomonas aeruginosa. The ISME Journal, 15(5), 1420–1433. https://doi.org/10.1038/s41396-020-00860-3

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. Journal of Veterinary Diagnostic Investigation, 21(4), 446–453. https://doi.org/10.1177/104063870902100405

Willms, I. M., Kamran, A., Aßmann, N. F., Krone, D., Bolz, S. H., Fiedler, F., & Nacke, H. (2019). Discovery of novel antibiotic resistance determinants in forest and grassland soil metagenomes. Frontiers in Microbiology, 10, 460. https://doi.org/10.3389/fmicb.2019.00460

Wood, R. L., Jensen, T., Wadsworth, C., Clement, M., Nagpal, P., & Pitt, W. G. (2020). Analysis of identification method for bacterial species and antibiotic resistance genes using optical data from dna oligomers. Frontiers in Microbiology, 11, 257. https://doi.org/10.3389/fmicb.2020.00257

Woods, E. C., & McBride, S. M. (2017). Regulation of antimicrobial resistance by extracytoplasmic function (ECF) sigma factors. Microbes and Infection, 19(4–5), 238–248. https://doi.org/10.1016/j.micinf.2017.01.007

Yang, M.-R., & Wu, Y.-W. (2022). Enhancing predictions of antimicrobial resistance of pathogens by expanding the potential resistance gene repertoire using a pan-genome-based feature selection approach. BMC Bioinformatics, 23(S4), 131. https://doi.org/10.1186/s12859-022-04666-2

Yoo, J.-S., Oh, G.-S., Ryoo, S., & Roe, J.-H. (2016). Induction of a stable sigma factor SigR by translation-inhibiting antibiotics confers resistance to antibiotics. Scientific Reports, 6(1), 28628. https://doi.org/10.1038/srep28628

Zeng, X., & Lin, J. (2013). Beta-lactamase induction and cell wall metabolism in Gram-negative bacteria. Frontiers in Microbiology, 4. https://doi.org/10.3389/fmicb.2013.00128

CHAPTER 5. PHENOTYPIC ANTIBIOTIC RESISTANCE PREDICTION IN *MANNHEIMIA HAEMOLYTICA* ISOLATES FROM CATTLE WITH BOVINE RESPIRATORY DISEASE USING RESISTANCE DATABASES AND MACHINE LEARNING

Carmen L. Wickware[1], Mohit Verma[2] and Timothy A. Johnson[1]
[1] Purdue University, Department of Animal Sciences, West Lafayette, IN 47907 USA
[2] Purdue University, Department of Agricultural and Biological Engineering, West Lafayette, IN

Acknowledgement: This work was supported by the Agriculture and Food Research Initiative Competitive Grants Program Award 2020-68014-31302 from the U.S. Department of Agriculture National Institute of Food and Agriculture. This work was also supported by a Purdue University Agricultural Science and Extension for Economic Development (AgSEED) grant. The funding organizations played no role in the design of the study; collection, analysis, and interpretation of data; or in writing the manuscript.

Abstract

Mannheimia haemolytica is one of the most common causative agents of bovine respiratory disease (BRD). Previous studies on phenotypic and genotypic antibiotic resistance in *M. haemolytica* have detailed the different relationships between virulence (e.g., outer membrane proteins) and antibiotic resistance through phenotype or genotype. Phenotypic resistance in *M. haemolytica* to macrolide and fluoroquinolone antibiotics is still fairly low, with studies reporting the majority of isolates below the intermediate breakpoint cutoff. However, isolate MIC values for these antibiotics have been steadily rising, suggesting a movement towards antibiotic resistance. Additionally, researchers have observed an influx of mobile genetic elements (MGE) such as integrative-conjugative elements (ICE) in the genomes of *M. haemolytica* previously found in other bacteria within the family *Pasteurellaceae*. These ICE commonly have around 100 different genes, with multiple resistance islands reported as conferring resistance to drugs in four antibiotic classes. Even so, the existence of the antibiotic resistance genes does not imply phenotypic resistance and BRD researchers have recently been studying phenotype and genotype in tandem to learn how the two resistances are interconnected. However, there are discrepancies in the accuracy of well-studied genotype prediction methods using antibiotic resistance databases. Prior investigations using resistance databases have reported high phenotype-genotype concordance for beta-lactam antibiotics ampicillin and penicillin (> 85%) but low for commonly used BRD

antibiotics such as tulathromycin and tilmicosin (<60%). The current study aimed to predict antibiotic resistance through the genome using machine learning (ML) by training reference-free models using known antibiotic resistance phenotypes coupled with sequence information. Comparison of genes found through antibiotic resistance database searches and sequences from the machine learning models showed an improvement for three of the seven antibiotics tested (danofloxacin, enrofloxacin, and tilmicosin) with 30% increase in accuracy for danofloxacin and enrofloxacin and a 10% increase for tilmicosin. There were four antibiotics (florfenicol, tetracycline, tildipirosin, and tulathromycin) for which a known ARG provided a more accurate prediction (>90% accuracy) than the machine learning models (>80% accuracy). When scanning the isolate genomes for potential annotations of the model output, various types of sequences were found. The annotations were coding sequences, such as DNA topoisomerase IV or *parA*, and non-coding sequences near tetracycline genes (*tetR, tetH*), MGE, or virulence genes. Using machine learning for antibiotics with low genotype-phenotype concordance in *M. haemolytica* helped improve the chances of accurately predicting antibiotic resistance. When tested on an external set of isolates for validation, the best predictor of antibiotic resistance (either ARG or ML) performed similarly to the test datasets for each antibiotic with the exception of tilmicosin which saw a 20% decrease in accuracy compared to model creation. Further research on the role of ML model sequences should be performed as they could potentially encode novel antibiotic resistance determinants.

5.1 Introduction

Mannheimia haemolytica (MH) is one of the main pathogens implicated in bovine respiratory disease (BRD). The bacteria are Gram negative, non-motile, non-spore forming, facultative anaerobic bacilli sometimes found in coccobacilli formations (Samanta & Bandyopadhyay, 2020). *M. haemolytica* is an exclusive animal pathogen originally classified as *Pasteurella haemolytica* due to its relation to the genus *Pasteurella* as well as the complete lysis of heme when grown on blood agar. However, after the identification of two separate biotypes of *P. haemolytica* (A and T), *P. haemolytica* biotype A (arabinose fermenting) was reclassified as *M. haemolytica*. *Mannheimia* was proposed as a novel genus by Angen et al (1999) due to differences in 16S rRNA gene structure compared to *Pasteurella spp.*, though the biotype designation A is still used in the identification of the > 20 serotypes of *M. haemolytica*.

Since BRD is an extremely costly disease to both the animal and producer, many management practices have been employed to improve the outcomes of animals by reducing risk factors. These practices include training calves on feed and water during weaning; weaning at least 45 days before transport from cow-calf operation to feedlot; castration and deworming of calves prior to transport from cow-calf operation to feedlot; and vaccinations (Hilton, 2015; Seeger et al., 2011; Theurer et al., 2021). Each of these practices focuses on a different type of risk factor: predisposing (e.g., transport stress and animal age), environmental (e.g., stocking density), or epidemiological (i.e., exposure to pathogenic microorganisms). By reducing these factors, animals ultimately have a lower risk of morbidity and mortality from BRD.

Treatment failure has been reported to affect 10-20% of cattle, with increased risk of mortality for each consecutive treatment (Cernicchiaro et al., 2013; Coetzee et al., 2019). Failed treatments are often the result of resistant bacteria. Antibiotic resistance (AR) can be intrinsic – the bacterial outer-membrane structure, efflux mechanisms, and epigenetic controls resulting in no antibiotic influx, activation, or target. AR can also be acquired through genetic mutations such as random insertions, deletions, and SNPs or horizontal gene transfer (HGT) mechanisms. HGT includes transduction (viral genetic transfer), natural transformation (bacterial engulfing of genetic material from the environment), and conjugation (bacteria-bacteria transfer through conjugative mechanisms).

While much information on the phenotypic or genotypic antibiotic resistance of MH isolates from BRD cases has been examined, there is a lack of investigations attempting to determine genotype-phenotype agreement. Previous analysis on phenotype to genotype comparisons in BRD pathogens has been done, but there is still a distinct gap in the amount of available data (Owen et al, 2017; Wickware and Johnson, not published). These previous studies have focused on analysis of three BRD pathogens, *Pasteurella multocida, Histophilus somni,* and *M. haemolytica*. In each of the studies, it was found there is low agreement of the phenotype and genotype in regard to antibiotic resistance. Where Owen et al (2017) found antibiotic resistance gene (ARG) database comparisons had little agreement with phenotype, Wickware and Johnson found that using machine learning improved this agreement. However, it was determined that comparing all three pathogens at once may not have actionable results – i.e., the resulting sequences may not be specific enough to an antibiotic to be used in a rapid assay.

For the reasons listed above, it was determined more analysis on the individual pathogens should be performed. The current study has utilized 119 isolates of MH from BRD cattle that have both antibiotic susceptibility testing (AST) and genome sequence data to establish phylogenetic relatedness, potential virulence, and antibiotic resistance patterns within the single species.

5.2 Materials and Methods

5.2.1 Isolates

Isolates of *Mannheimia haemolytica* (n = 92) were obtained from previously collected nasopharyngeal and deep lung swabs of cattle processed at the Purdue University Animal Disease Diagnostic Laboratory (PU; n = 24), Kansas State University Veterinary Diagnostic Laboratory (KSU; n = 49), and Texas A&M University Veterinary Medical Diagnostic Laboratory (TAMU; n = 19). Six PU isolates were from a previous experiment and were prepared for Illumina MiSeq using the TruSeq DNA PCR-free Library Prep kit (Illumina, San Diego, USA). Sequencing for these six isolates was performed at the Purdue Genomics Core using MiSeq sequencing (Illumina; 2 x 300 cycles). The remaining PU isolates as well as KSU and TAMU isolates were sent to KSU Veterinary Diagnostic Laboratory for sequence preparation using Illumina DNA Prep kit (Illumina, San Diego, USA) and sequenced using Illumina MiSeq (2 x 300 cycles).

5.2.2 Genome Assembly

Isolates of *M. haemolytica* (n = 26) previously sequenced by Owen et al (OA; 2017) were used in the current study. Sequence reads from OA isolates were downloaded from the NCBI Sequence Read Archive (Bioproject PRJNA306895). Sequence reads for all OA, PU, KSU, and TAMU isolates were trimmed for quality using Trimmomatic (v 0.39; Bolger et al., 2014) and sequence quality was assessed with fastQC (v 0.11.9). Quality checked sequence reads were assembled using SPAdes (v 3.13.0; Bankevich et al., 2012; Nurk et al., 2013) with default parameters for assembly with the exception of the *"--careful"* parameter to reduce the number of mismatches and short contigs.

Assembled reads were filtered to remove all contigs < 500bp. . Assembly quality was assessed using quast (v 3.2 ; Gurevich et al., 2013; Mikheenko et al., 2016) with default parameters.

Recorded statistics include number of contigs, max contig, total length, and N50. Full assembly statistics for PU, KSU, and TAMU isolates can be found in Appendix C.1.

5.2.3 Antibiotic susceptibility testing (AST) and ARG annotation

MIC values and interpretations for PU, KSU, and TAMU isolates are listed in Appendix C.2. and are published in Owen et al., (TableS1; 2017) for OA isolates. Due to variability in the AST plates used by different diagnostic labs, not all tested antibiotics have the same number of isolates.

ARGs were annotated using CARD-RGI (Alcock et al., 2019) with the following parameters:

--input_type contig -d wgs --local --exclude_nudge –clean

The "loose" matches were excluded, leaving only "strict" and "perfect" hits, to have the most confidence in the gene calls from CARD. Concordance was calculated for each antibiotic by taking the sum of the isolates with matching genotype (G) and phenotype (P) of antibiotic susceptible (S) or resistant (R) isolates (GR+PR and GS+PS) and dividing by the total sum of isolates (i.e., GR/PR and GS/PR and GR/PS and GS/PS) (Table 1).

5.2.4 Antibiotic resistance prediction using machine learning

Assembled genomes (n = 96) were used to create a table of k-mers (k=31) representing the presence or absence of the 31 base-long sequence in each genome. K-mer length 31 was used as it produced the best models in a previous study (see Chapter 4). The k-mers were created using Suffixerator and counted using Tallymer using GenomeTools (v 1.5.9; Kurtz et al., 2008; Gremme et al., 2013). A metadata file was created for each antibiotic that contained the assembly names and whether the isolates were phenotypically resistant (1) or susceptible (0) to a particular antibiotic. To create a model of AR prediction, more than 10 isolates in the resistant or susceptible groups were required (Drouin et al., 2016). The k-mer table and metadata were used to create a dataset for KOVER. The dataset for each antibiotic was used to predict resistance in KOVER with the Set Covering Machine algorithm (Drouin et al., 2016, 2019; Marchand & Shawe-Taylor, 2002). Briefly, the dataset was split into training and testing groups (66 and 33% of the data, respectively). A model was learned for each antibiotic using the training data and validated using

the testing data. The output was a set of rules (31-mers) each assigned an importance based on how often the rule was considered for the model (Drouin et al., 2016, 2019) and the equivalent rules (ER) – the set of 31-mers considered equally important for predicting AR in the model. Model statistics including error, F1-score (harmonic mean), specificity, sensitivity, and precision for both training and testing groups were also included as output for each antibiotic.

(Drouin, 2018). First, AR models including ER were used in a BLAST search against the family *Pasteurellaceae* (NCBI:txid712). Alignments to NCBI reference genomes were not always high quality. The best alignments (as determined through e-value, length, and percent coverage of query to subject) to the reference sequences were called reference alignments. Low e-values ($e < 10^{-50}$) and total query length at 100% coverage were considered as the reference alignment. In a case where the e-value was higher than 10^{-50} and/or coverage was less than 100%, the best alignment possible was considered. Sequence query information can be found in Appendix C.3 (seq info). BLAST results identified as the reference alignment were examined to identify genes or regions of interest in the proximity of the reference alignment.

Second, the reference alignments were downloaded from NCBI as GenBank files so the ER could be aligned to the reference sequence using UGENE (v 40.1; Okonechnikov et al., 2012). The UGENE function "Find Pattern" was used to align the model rule sets to the matching reference alignment (i.e., model 1 for ampicillin with GenBank file from model 1 ER). Through this process a filtered alignment file was saved in fasta format for use in later steps.

5.2.5 Model alignment to genomes

The following procedures, used in a prior study (see Chapter 4), were modified from the suggested steps outlined by Drouin to further characterize the output of Kover model analysis Annotation and Target selection

Genomes were annotated using Rapid Annotation and Subsystem Technology (Aziz et al., 2008; Brettin et al., 2015; Overbeek et al., 2013). Using the aligned rulesets from UGENE, a BLAST search was performed within RAST with the alignments as the query and the isolate genomes as the subject. For each model, if the model type was "presence", either conjunction or disjunction, the genomes considered to be in the resistant group were the subject. If the model type was "absence", either conjunction or disjunction, the genomes considered to be sensitive were the

subject. Resulting output was searched not only to determine the sequence location within the genomes, but to examine the surrounding areas on the contigs for genes of interest.

5.2.6 Final validation

19 isolates (14 KSU, 2 TAMU, 3 PU) were randomly chosen using random number generation and removed from the testing data set to be used for final validation of the models. The isolates represent ~15% of the total isolates used in the study. The AST data for the 19 isolates were removed from the model creation data by a researcher not involved in the study to keep the authors blind to the antibiotic resistance phenotype. Validation was performed by determining the resistance using the best predictor (i.e., ML model or ARG; table 2) for each antibiotic as the query in a BLAST search. The predictions were then compared to the phenotype AST results and summarized as number incorrectly predicted out of 19 by the same researcher who removed the AST data.

5.2.7 Data availability

Data from Owen et al is available through their published work (Owen et al., 2017). The data used from Owen et al in this study can be found in supplementary file XX. Sequence reads from PU, KSU, and TAMU isolates can be found at NCBI SRA and assembled genomes are available in the NCBI Genome Repository (BioProject XXXXXXXXX). Bioinformatics scripts, input, and output files for Kover and CARD available at github.com/clwickwa/BovineRespiratoryDisease/Mannheimia/scripts.

Table 5.1 Concordance between antibiotic resistance phenotype and genes. P: phenotype; G: genotype; R: resistant; S: susceptible; ML: machine learning

		P:R		P:S						
		G:R	G:S	G:R	G:S	Precision	Sensitivity/Recall	Specificity	F1	Accuracy
ML antibiotics	Danofloxacin	0	38	1	59	0.00	0.00	0.98	N/A	0.60
	Enrofloxacin	0	38	1	59	0.00	0.00	0.98	N/A	0.60
	Florfenicol	25	6	4	63	0.86	0.81	0.94	0.83	0.90
	Tetracycline	32	2	0	32	1.00	0.94	1.00	0.97	0.97
	Tildipirosin	15	1	1	26	0.94	0.94	0.96	0.94	0.95
	Tilmicosin	31	26	5	36	0.86	0.54	0.88	0.67	0.68
	Tulathromycin	24	1	1	49	0.96	0.96	0.98	0.96	0.97
Non-ML antibiotics	Ampicillin	37	2	14	45	0.73	0.95	0.76	0.82	0.84
	Ceftiofur	1	0	50	47	0.02	1.00	0.48	0.04	0.49
	Clindamycin	32	47	0	0	1.00	0.41	N/A	0.58	0.41
	Gamithromycin	16	1	0	26	1.00	0.94	1.00	0.97	0.98
	Gentamicin	26	0	27	20	0.49	1.00	0.43	0.66	0.63
	Neomycin	38	0	2	17	0.95	1.00	0.89	0.97	0.96
	Oxytetracycline	12	0	0	11	1.00	1.00	1.00	1.00	1.00
	Penicillin	31	3	1	20	0.97	0.91	0.95	0.94	0.93
	Spectinomycin	56	10	3	29	0.95	0.85	0.91	0.90	0.87
	Sulfadimethoxine	44	11	6	37	0.88	0.80	0.86	0.84	0.83
	Tiamulin	0	59	0	29	N/A	0.00	1.00	N/A	0.33
	Trimeth/sulfa	3	2	36	22	0.08	0.60	0.38	0.14	0.40
	Tylosin tartrate	35	24	1	38	0.97	0.59	0.97	0.74	0.74

Table 5.2 Comparison of F1 scores, accuracy, and error between resistance database query (ARG) and machine learning (ML) for each tested antibiotic

	ARG			ML		
Antibiotic	F1	Accuracy	Error	F1	Accuracy	Error
Danofloxacin	N/A	0.60	0.40	0.8	0.875	0.125
Enrofloxacin	N/A	0.60	0.40	0.9	0.938	0.062
Florfenicol	0.83	0.90	0.10	0.67	0.88	0.12
Tetracycline	0.97	0.97	0.03	0.87	0.917	0.083
Tildipirosin	0.94	0.95	0.05	0.75	0.88	0.12
Tilmicosin	0.67	0.68	0.32	0.8	0.78	0.22
Tulathromycin	0.96	0.97	0.03	0.78	0.84	0.16

5.3 Results and Discussion

5.3.1 Sequence read and assembly quality

The quality of the sequence reads varied between different isolate origins. OA isolates had the shortest length and the lowest quality compared to the isolates sequenced at PU and KSU. After assembly, the quality of OA isolate sequences improved compared to the others but were still lower overall (Appendix C.1). However, the quality of the OA assemblies compared to those previously generated by OA was only marginally improved (Appendix C.1; Owen et al, 2017).

Table 5.3 KOVER model for breakpoint MIC. Model type: presence (P) or absence (A) of the ruleset predicts resistance; Importance: weighted value for rule – designates how often the rule was found in the model; For multiple rulesets - Conjunction: all rulesets pre predict resistance (logical AND) or Disjunction: one ruleset needs to be present or absent to predict resistance (logical OR).

Antiboitic	Model type (importance)		# equiv rules	Annotation
Danofloxacin	Disjunction	P (0.69)	7	Sequence near *tetR*
		P(0.52)	31	Sequence upstream of toxin-antitoxin genes *higA/higB*
		A(0.17)	35	Sequence upstream of IS481 transposase
		P(0.48)	31	Section of DNA topoisomerase IV
Enrofloxacin	Disjunction	P (0.70)	7	Sequence associated with *tetR*
		P (0.48)	31	Sequence upstream of toxin-antitoxin genes *higA/higB*
		P (0.15)	137	Five regions associated with virulence and survival
Florfenicol	Disjunction	P (0.76)	7	Sequence near *tetR*
		P(0.45)	55	Sequence *tetR*
Tetracycline	Conjunction	P (0.96)	31	Duplicated sequence between hypothetical protein CDS and exodeoxyribonuclease I
		A (0.50)	2	Multiple copies of ISSod 13/IS1595 transposase
Tildipirosin	Disjunction	P (0.58)	251	Five regions with AR and MGE sequences
		P (0.50)	31	Hypothetical protein CDS
Tilmicosin	Conjunction	A (0.74)	1049	ParA/ParB
		A (0.58)	26	Phage tail tip – host specificity J
Tulathromycin	Disjunction	P (0.76)	7	Sequence associated with *tetR*
		A (0.38)	1	Unspecific sequence matches
		P (0.38)	23	Upstream and start of Laccase

5.3.2 AST phenotype and potential ARGs

The majority of antibiotics tested by broth microdilution had more than 30% of the isolates with phenotypic resistance (Appendix C.2). Tulathromycin and trimethoprim+sulfadimethoxazole (TMS) had 13% and 10% of isolates with resistance, spectinomycin had 9%, and ceftiofur only had 1% of isolates with phenotypic resistance. Phenotypic resistance to tulathromycin and spectinomycin has been seen in BRD-associated pathogens, including *M. haemolytica,* in previous studies (Holschbach et al., 2020; Owen et al., 2017; Stanford et al., 2020; Torres-Blas et al., 2021). Trimethoprim/sulfamethoxazole is rarely used for treatment of BRD though resistance has been

observed in *M. haemolytica* previously with similar levels of phenotypic resistance to the current study (Clinical and Laboratory Standards Institute, 2020; Owen et al., 2017). Ceftiofur, a third generation cephalosporin, has been used in the treatment of BRD due to its broad-spectrum, bactericidal action and low incidence of resistance in respiratory pathogens (Owen et al., 2017; Torres-Blas et al., 2021).

In addition to differing numbers of observed phenotypically antibiotic resistant isolates, there were varying antibiotic resistant genes found (Appendix C.3). The isolates from each origin (KSU, OA, PA, TAMU) had some shared ARGs. However, no resistance gene was found in more than 53% of the isolates with the exception of the disinfectant resistance gene *qacG* (98%) and tetracycline resistance gene *tetH* (62%) (Appendix C.3). Unsurprising was the fact that isolates from the same origin seemed to share a similar ARG profile with isolates from the same origin.

Antibiotic resistance genes identified in the current study include those conferring resistance to all of the antibiotic classes used in the treatment of BRD – bla_{ROB-1}, bla_{OXA-2}, $bla_{TEM-229}$, bla_{ROB-2}, bla_{CMY-2}, and bla_{CRP} (beta-lactam); *arlR, arlS, emrR, emrA, emrB, Saur_norA, norC,* and *sdrM* (fluoroquinolone); *erm(42), erm(47), mphE,* and *msrE* (macrolide-lincosamide-stretogramin; MLS); *floR* and *Ccol_ACT_CHL* (phenicol); and *tetA, tetH, tetT, tet38, mepA,* and *mepR* (tetracycline) – as well as many others (Appendix C.3).

5.3.3 ARG prediction and machine learning model statistics

Concordance between ARGs and phenotypic resistance was above 90% for florfenicol, tildipirosin, tetracycline, and tulathromycin (Table 5.1). However, danofloxacin, enrofloxacin, and tilmicosin had less than 70% agreement with the phenotype AST. Though roughly 40% of the isolates had phenotypic resistance to danofloxacin and enrofloxacin, fluoroquinolone resistance genes were only found in one isolate, MH084, which was phenotypically susceptible to danofloxacin and enrofloxacin (Appendix C.2). The lack of fluoroquinolone resistance genes led to low accuracy and harmonic mean for the phenotypic danofloxacin and enrofloxacin resistance when using ARG as the predictor.

Additionally, using ARGs for macrolide antibiotics resulted in low sensitivity for predicting tilmicosin resistance, where half the resistant isolates were misclassified as susceptible (Table 5.1). Interestingly, machine learning improved the accuracy and harmonic mean compared to use of known antibiotic resistance genes for three of the seven tested antibiotics. While the

accuracy between machine learning models and phenotype AST was greater than 75% for all tested antibiotics, the harmonic mean for florfenicol was less than 75% (Table 5.4).

Antibiotic resistance genes were the most accurate predictors of resistance for florfenicol, tetracycline, tildipirosin, and tulathromycin while danofloxacin, enrofloxacin, and tilmicosin resistance were more accurately predicted from their respective machine learning models. There was an inverse relationship between ARG accuracy and ML accuracy. In the four antibiotics with $\geq$ 90% prediction accuracy using ARGs, the ML accuracy was < 90% whereas in the three antibiotics with between 60-70% accuracy from ARG prediction, the ML accuracy improved at least 10%. The same pattern was seen with the harmonic mean (Table 5.2). Additionally, tetracycline had the highest harmonic mean of the antibiotics with 97% accuracy using ARGs. Tetracycline resistance prediction was more accurate using ARGs (97% concordance) compared to the ML model (92% concordance; Table 5.2).

5.3.4 Machine learning model alignment and annotation

Each created model for the seven tested antibiotics resulted in multiple rules (Table 5.). The rules for danofloxacin, enrofloxacin, florfenicol, tildipirosin, and tulathromycin were part of disjunctive models (i.e., the presence/absence of any one rule must be satisfied) while tetracycline and tilmicosin rules were conjunctive (i.e., all rules must be satisfied). The summary of this information can be found in Table 5.3. Though the ARG models were more accurate for florfenicol, tetracycline, tildipirosin, and tulathromycin, the ML models were still analyzed to determine genetic context for why the models did not perform well.

Four antibiotics shared the same first rule – danofloxacin, enrofloxacin, florfenicol, and tulathromycin (Appendix C.3 and C.4). This rule was 37 base long (7 ER) and matched a small, non-coding sequence found upstream of the tetracycline repressor gene *tetR* (Appendix C.3 and C.4). The sequence matched 97% of the phenotypically resistant isolates for each antibiotic. However, there were some isolates considered phenotypically susceptible to one antibiotic but resistant to another, decreasing the importance of the rule due to phenotypically resistant isolates without the matching rule sequence (Appendix C.4).

In addition to the shared rule for the four antibiotics above, danofloxacin and enrofloxacin shared the same second rule, the presence of which was predictive for resistance of these two antibiotics. This is less surprising than the outcome above as both danofloxacin and enrofloxacin

are both fluoroquinolone antibiotics. Rule 2 for danofloxacin and enrofloxacin was 61 bases long and matched a non-coding sequence associated with one part of a toxin-antitoxin system, *higA,* a mechanism for virulence gene regulation (Appendix C.4; Wood and Wood, 2016). Rule 2 matched 97% of the resistant isolates and was found in susceptible isolates.

Table 5.4 Kover model metrics for ML tested antibiotics

	Training Data							Testing Data						
Antiboitic	Susc	Res	Error	F1	Sens/Recall	Spec	Precision	Susc	Res	Error	F1	Sens/Recall	Spec	Precision
Danofloxacin	37	29	0.0303	0.96	0.97	0.97	0.96	23	9	0.13	0.8	0.89	0.87	0.73
Enrofloxacin	37	29	0.061	0.93	0.9	0.97	0.96	23	9	0.06	0.9	1	0.91	0.82
Florfenicol	40	26	0.075	0.91	0.96	0.9	0.86	27	5	0.12	0.67	0.8	0.89	0.57
Tetracycline	21	27	0.1	0.9	0.81	1	1	15	9	0.08	0.87	0.78	1	1
Tildipirosin	21	12	0.0303	0.96	0.92	1	1	12	4	0.12	0.75	0.75	0.92	0.75
Tilmicosin	23	43	0.061	0.96	1	0.83	0.91	17	15	0.22	0.8	0.93	0.65	0.7
Tulathromycin	32	18	0.06	0.92	1	0.91	0.86	18	7	0.16	0.78	1	0.78	0.64

However, while the majority of isolates with phenotypic resistance to fluoroquinolones had 100% sequence identity to the rule two sequence, all the isolates with phenotypic susceptibility had a SNP in the middle of the match (C to T; Appendix C.4). Of the nine phenotypically resistant isolates with the same SNP, only one was misclassified by the danofloxacin and enrofloxacin models (Appendix C.4).

The two additional rules for danofloxacin resistance and the one additional rule for enrofloxacin resistance were not shared. Danofloxacin rule 3 was 81 bases long and matched a non-coding sequence (nCDS) just upstream of an IS481-like transposase and was surrounded by MGE genes – integrases, plasmid stabilization proteins, and unclassified mobile genetic elements (Appendix C.4). The absence of this gene was considered a predictor for danofloxacin resistance and 88% of isolates with phenotypic susceptibility to danofloxacin had multiple copies of this sequence throughout the genome, 9% had only one copy, and only one isolate, MH084, did not have any matches (Appendix C.4). Rule 4 for danofloxacin resistance, for which the presence predicted resistance, was a 61 base long sequence with 100% of phenotypically resistant isolates having the sequence though some had only 83% sequence similarity (nine isolates) with the filtered alignment (Appendix C.3). This rule matched a section of DNA topoisomerase, one of the targets of danofloxacin. The genes surrounding the rule sequence have been identified as virulence genes in other studies of MH, sialic acid metabolism and capsule formation (Appendix C.4; Cai et al., 2020; Samanta and Bandyopadhyay, 2020). In eight of the nine isolates with the lower sequence similarity to the filtered alignment, there was a SNP (A to G) in the middle of the sequence. There was one isolate, MH012, with 83% similarity to the sequence that had several mismatches, a high e-value (e -5), and completely different surrounding genes indicating this rule does not help predict resistance in the isolate.

Enrofloxacin rule 3 was split into seven different regions, all within the same section of the isolate genomes. Each region was between 40-107 bases long, had a sequence similarity between the isolate genomes and the filtered alignment between 79 and 100%, and was found in all isolates with phenotypic enrofloxacin resistance (Appendix C.3). The annotation of the seven regions of the genome that aligned with rule 3 included genes useful during survival and invasion of macrophages, one of the main modes of pathogenesis for MH (Appendix C.4; Visick et al., 1998; Mathur et al., 2006).

Misclassification of isolates was seen for all antibiotics. The ML model for tilmicosin resulted in the most misclassified isolates and all but one were phenotypically susceptible to tilmicosin (Table 3). In contrast, tildipirosin had one misclassified isolate, MH084. For models with shared rules (danofloxacin, enrofloxacin, florfenicol and tulathromycin), typically only those isolates that were either unique or missing from one of the resistant datasets that were misclassified (Appendix C.5).

Though the ML models for florfenicol, tetracycline, tildipirosin, and tulathromycin were not as accurate as their ARG predictions (Table 5.2), it is still important to look at the results of the models to determine why this could be the case. The second florfenicol rule partially matched the same region as florfenicol rule 1, mentioned previously as the shared first rule for four antibiotics. However, the rule was split into three regions, one 49 bases long and matching a section of the *tetH* gene and two different regions upstream of *tetR* that were 54 and 61 bases long (Appendix C.4). When looking at the differences between rule 1 and rule 2, it seems the regions in rule two have larger sections of the sequences that do not match those of rule 1 suggesting different alleles of the gene (Appendix C.3).

The model for tetracycline consisted of two rules and was a conjunction model with the presence of rule 1 and the absence of rule 2 both needing to be satisfied for resistance (Table 5.3). Tetracycline rule 1, a 61 base long sequence matching 91% of the isolates. This sequence was duplicated in all isolates with matches with the sequence being located between a CDS for a hypothetical protein and another hypothetical protein CDS for one duplicate or exodeoxyribonuclease I in the other duplicate sequence (Appendix C.4). The sequences surrounding rule 1 were different genes annotated as DNA repair, sodium/amino acid symport, phage component, and iron acquisition. Rule 2 was split into two regions, both associated with transposases, though one region was 684 bases long and matched ISSod 13 and the other was 436 bases long and matched IS1595 (Appendix C.3 and C.4). Multiple copies of these IS were found with 91% sequence similarity to the filtered alignment and located on contigs with 23S methyltransferases and various membrane permease genes (Appendix C.4).

The tildipirosin model was a disjunction model and had two rules with either the presence of rule 1 or rule 2 predicting resistance (Appendix C.4). Rule 1 was split into five regions all of which were found within the same section of one contig. Region 1 was annotated as an iron transport permease gene, regions 2 and 3 were different sequences associated with *tetR* and *tetH*

(i.e., between the two genes, upstream of *tetR*, or downstream of *tetH*), region 4 was downstream of the same *tetH* gene with part of the sequence annotated as the beginning of an ISL3 transposase, and region 5 being located at the beginning or end of the contig depending on the strand (Appendix C.4). Other CDS on these contigs were annotated as various genes related to MH virulence and pathogenesis – iron transport, RTX toxin genes, macrophage infectivity genes, and cell division machinery. The second rule was found as the only CDS on the contig in each of the genomes with matching sequences (Appendix C.3 and C.4).

As discussed previously, tulathromycin had the same rule 1 as danofloxacin, enrofloxacin, and florfenicol. The disjunctive model for tulathromycin consisted of three rules where either the presence of rule 1, absence of rule 2, or presence of rule 3 needed to be satisfied for resistance to tulathromycin. Rule 3, a 53 base long sequence, was found in 96% of isolates with phenotypic resistance to tulathromycin (Appendix C.3). Again, isolate MH084 was phenotypically resistant to tulathromycin but the ruleset was not found within the genome (Appendix C.4). Rule 3 was annotated as the beginning of laccase as well as a section upstream of laccase with surrounding genes related to stress response (Appendix C.4). The second rule of tulathromycin consisted of a single equivalent rule (Table 3) and was not able to be aligned as it matched too many varying sections of the genome, perhaps lending to the poor performance of the tulathromycin model in predicting phenotypic resistance (Appendix C.3).

Several isolates were found as testing and training errors for more than one ML model (Appendix C.5). In particular, MH019 and MH084 were found as errors in four different antibiotics and MH082 was an error in five of the seven antibiotics. Of these three isolates, MH082 and MH084 were consistently found with different gene patterns when examining annotations of the genes surrounding the model rules in each isolate. Isolate MH084 (PA) also had a genome twice the expected size at 4.79 Mb. Matrix assisted laser desorption-ionization, time-of-flight mass spectrometry (MALDI-TOF MS) identification of MH084 resulted in the isolate being considered *M. haemolytica*. However, it is possible that the isolate was misidentified. Misidentification of uncommon pathogens by MALDI-TOF MS has been seen in human clinical studies and the accuracy of MALDI-TOF MS in identifying BRD pathogens has been reported as less than 80% in pure and mixed cultures (AbdulWahab et al., 2015; Puchalski et al., 2016; Van Driessche et al., 2019).

5.3.5 Performance of models on new data

Using the most accurate prediction method based on the model creation dataset resulted in similar accuracy of resistance prediction in the final validation within 3% for enrofloxacin, florfenicol, tetracycline, and tildipirosin (Table 5.5). Danofloxacin and tulathromycin resistance prediction had lower error than their model predictions with danofloxacin (ML model) having 7% lower error and tulathromycin (ARG prediction) with 6% lower error (Table 4). Tilmicosin had 25% higher error than the machine learning model. As tilmicosin is a commonly used BRD antibiotic (trade name Micotil), it is imperative to determine a suitable predictive model.

5.3.6 Future considerations

In the current study, phenotypic identification of isolates was limited to one species. However, it is important to note that serotype plays a major role in the pathogenesis of the isolates (Fodor et al., 2010; Hauglund et al., 2015; Klima et al., 2017). Serotype clustering from genome and ID with molecular methods should be performed to determine any connection between antibiotic resistance and serotype.

For predictions of antibiotic resistance that were better using ARGs, the use of random forest to consider which ARG is predictive of phenotype for each drug class would be useful information for the creation of an assay from the sequence data (Kouchaki et al., 2020; Moradigaravand et al., 2018). The current study found multiple ARGs across isolates for each antibiotic class, but to determine which ARGs are the most important predictors, further analysis is warranted.

To use the results of the current study in the determination of the best treatment for animals with BRD, it is assumed that animal and/or farm information (prior treatment, animal age, farm disease data, season, etc.) should be included in the decision-making process, as each of these factors can increase the risk of treatment failure. In addition to these factors, the pathogen abundance and microbial community patterns the upper respiratory tract may also aid in treatment decisions (Centeno-Martinez et al., 2022)

Table 5.5 Final validation results

	Antibiotic	Incorrect Predictions	Percent error	Prediction Model Error
ML prediction	Danofloxacin	1	5%	12.5%
	Enrofloxacin	1	5%	6.2%
	Tilmicosin	9	47%	22%
ARG prediction	Florfenicol	3	15%	12%
	Tetracycline	2	10%	8.3%
	Tildipirosin*	1	10%	12%
	Tulathromycin	2	10%	16%

* 9 of the 19 isolates did not have AST data for tildipirosin; final validation is for total of 10 isolates

5.3.7 Future considerations for the treatment of animals with ARB

To determine how HGT plays a role in the gain of antibiotic resistances within MH, researchers have studied the genomes of MH isolates. Mobile genetic elements (MGE) such as transposons and integrative-conjugative elements (ICE) allow transfer of genetic materials between bacteria and have been identified in MH isolates from cattle with BRD. ICE donor and recipient cells typically need to be from the same taxonomic family and an ICE with identical backbones to those found in *P. multocida* (*ICEPmu1*) have been found in MH isolates (Eidam et al., 2015; Owen et al., 2017). Clawson et al. (2016) examined MH isolates to determine genotypes related to pathogenesis and found *ICEMh1* sequences were more frequently found in isolates from lungs than nasopharynx or nasal samples. *ICEMh1* was reported as having up to 107 total genes and resistance genes for aminoglycosides, tetracyclines, and sulfonamides (Eidam et al., 2015).

5.4 Conclusions

Bovine respiratory disease is a major concern to both dairy and beef cattle producers, making the study of the causative agents and antibiotic resistance vital. Through the current study, we were able to determine sequences that improved the prediction of resistance to antibiotics commonly used in the treatment of BRD in the United States. PCR amplification targeting known resistance genes is still the most accurate method for four of the seven antibiotics. However, for the fluoroquinolone antibiotics and tilmicosin, machine learning models targeting other sequences were able to improve phenotypic resistance prediction accuracy by 10% for tilmicosin and 20-30% for the fluroquinolones.

Antibiotic resistance prediction in commensal and pathogenic, animal-associated bacteria is an important area of study. The treatment of BRD is highly error-prone due to the necessary use of empiric antibiotic treatment – treatment that is given prior to the determination of etiology, source of infection, or resistance profile of the disease-causing pathogen. Empiric treatments allow for an immediate attempt at treating the animal rather than waiting days for laboratory confirmed information. However, empiric treatment can result in failed treatments, i.e., use of ineffective antibiotics that result in fatality require retreating animals with different antibiotics.

5.5 References

AbdulWahab, A., Taj-Aldeen, S., Ibrahim, E. B., Talaq, E., Abu-Madi, M., & Fotedar, R. (2015). Discrepancy in MALDI-TOF MS identification of uncommon Gram-negative bacteria from lower respiratory secretions in patients with cystic fibrosis. Infection and Drug Resistance, 83. https://doi.org/10.2147/IDR.S80341

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L. V., Cheng, A. A., Liu, S., Min, S. Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R. E., Nasir, J. A., Oloni, M., Speicher, D. J., Florescu, A., Singh, B., … McArthur, A. G. (2019). CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Research, gkz935. https://doi.org/10.1093/nar/gkz935

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008). The RAST server: rapid annotations using subsystems technology. BMC Genomics, 9(1), 75. https://doi.org/10.1186/1471-2164-9-75

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology, 19(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics, 30(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Scientific Reports, 5(1), 8365. https://doi.org/10.1038/srep08365

Cai, Y., Varasteh, S., van Putten, J. P. M., Folkerts, G., & Braber, S. (2020). Mannheimia haemolytica and lipopolysaccharide induce airway epithelial inflammatory responses in an extensively developed ex vivo calf model. Scientific Reports, 10(1), 13042. https://doi.org/10.1038/s41598-020-69982-0

Centeno-Martinez, R. E., Glidden, N., Mohan, S., Davidson, J. L., Fernández-Juricic, E., Boerman, J. P., Schoonmaker, J., Pillai, D., Koziol, J., Ault, A., Verma, M. S., & Johnson, T. A. (2022). Identification of bovine respiratory disease through the nasal microbiome. Animal Microbiome, 4(1), 15. https://doi.org/10.1186/s42523-022-00167-y

Cernicchiaro, N., White, B. J., Renter, D. G., & Babcock, A. H. (2013). Evaluation of economic and performance outcomes associated with the number of treatments after an initial diagnosis of bovine respiratory disease in commercial feeder cattle. 74(2), 10.

Clawson, M. L., Murray, R. W., Sweeney, M. T., Apley, M. D., DeDonder, K. D., Capik, S. F., Larson, R. L., Lubbers, B. V., White, B. J., Kalbfleisch, T. S., Schuller, G., Dickey, A. M., Harhay, G. P., Heaton, M. P., Chitko-McKown, C. G., Brichta-Harhay, D. M., Bono, J. L., & Smith, T. P. L. (2016). Genomic signatures of Mannheimia haemolytica that associate with the lungs of cattle with respiratory disease, an integrative conjugative element, and antibiotic resistance genes. BMC Genomics, 17(1), 982. https://doi.org/10.1186/s12864-016-3316-8

Clinical and Laboratory Standards Institute. (2020). Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals (5th ed.). CLSI.

Coetzee, J. F., Magstadt, D. R., Sidhu, P. K., Follett, L., Schuler, A. M., Krull, A. C., Cooper, V. L., Engelken, T. J., Kleinhenz, M. D., & O'Connor, A. M. (2019). Association between antimicrobial drug class for treatment and retreatment of bovine respiratory disease (BRD) and frequency of resistant BRD pathogen isolation from veterinary diagnostic laboratory samples. PLOS ONE, 14(12), e0219104. https://doi.org/10.1371/journal.pone.0219104

Drouin, A. (2018, August). Interpreting Kover models. Kover Documentation. https://aldro61.github.io/kover/doc_interp.html

Drouin, A., Giguère, S., Déraspe, M., Marchand, M., Tyers, M., Loo, V. G., Bourgault, A.-M., Laviolette, F., & Corbeil, J. (2016). Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons. BMC Genomics, 17(1), 754. https://doi.org/10.1186/s12864-016-2889-6

Drouin, A., Letarte, G., Raymond, F., Marchand, M., Corbeil, J., & Laviolette, F. (2019). Interpretable genotype-to-phenotype classifiers with performance guarantees. Scientific Reports, 9(1), 4071. https://doi.org/10.1038/s41598-019-40561-2

Eidam, C., Poehlein, A., Leimbach, A., Michael, G. B., Kadlec, K., Liesegang, H., Daniel, R., Sweeney, M. T., Murray, R. W., Watts, J. L., & Schwarz, S. (2015). Analysis and comparative genomics of ICEMh1, a novel integrative and conjugative element (ICE) of Mannheimia haemolytica. 5. https://doi.org/doi:10.1093/jac/dku361

Fodor, L., Varga, J., Hajtós, I., & Szemerédi, Gy. (2010). Serotypes of Pasteurella haemolytica isolated from sheep, goats and calves. Zentralblatt Für Veterinärmedizin Reihe B, 31(1–10), 466–469. https://doi.org/10.1111/j.1439-0450.1984.tb01324.x

Gremme, G., Steinbiss, S., & Kurtz, S. (2013). GenomeTools: A comprehensive software library for efficient processing of structured genome annotations. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 10(3), 645–656. https://doi.org/10.1109/TCBB.2013.68

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. Bioinformatics, 29(8), 1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Hauglund, M. J., Tatum, F. M., Bayles, D. O., Maheswaran, S. K., & Briggs, R. E. (2015). Genome Sequences of serotype A6 Mannheimia haemolytica isolates D174 and D38 recovered from bovine pneumonia. Genome Announcements, 3(2), e00086-15. https://doi.org/10.1128/genomeA.00086-15

Hilton, W. M. (2015). Management of preconditioned calves and impacts of preconditioning. Veterinary Clinics of North America: Food Animal Practice, 31(2), 197–207. https://doi.org/10.1016/j.cvfa.2015.03.002

Holschbach, C. L., Aulik, N., Poulsen, K., & Ollivett, T. L. (2020). Prevalence and temporal trends in antimicrobial resistance of bovine respiratory disease pathogen isolates submitted to the Wisconsin Veterinary Diagnostic Laboratory: 2008–2017. Journal of Dairy Science, 103(10), 9464–9472. https://doi.org/10.3168/jds.2019-17570

Klima, C. L., Zaheer, R., Briggs, R. E., & McAllister, T. A. (2017). A multiplex PCR assay for molecular capsular serotyping of Mannheimia haemolytica serotypes 1, 2, and 6. Journal of Microbiological Methods, 139, 155–160. https://doi.org/10.1016/j.mimet.2017.05.010

Kouchaki, S., Yang, Y., Lachapelle, A., Walker, T. M., Walker, A. S., CRyPTIC Consortium, Peto, T. E. A., Crook, D. W., & Clifton, D. A. (2020). Multi-label random forest model for tuberculosis drug resistance classification and mutation ranking. Frontiers in Microbiology, 11, 667. https://doi.org/10.3389/fmicb.2020.00667

Kurtz, S., Narechania, A., Stein, J. C., & Ware, D. (2008). A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes. BMC Genomics, 9(1), 517. https://doi.org/10.1186/1471-2164-9-517

Marchand, M., & Shawe-Taylor, J. (2002). The set covering machine. Journal of Machine Learning Research, 3, 723–746.

Mathur, D., Malik, G., & Garg, L. C. (2006). Biochemical and functional characterization of triosephosphate isomerase from Mycobacterium tuberculosis H37Rv. FEMS Microbiology Letters, 263(2), 229–235. https://doi.org/10.1111/j.1574-6968.2006.00420.x

Mikheenko, A., Saveliev, V., & Gurevich, A. (2016). MetaQUAST: Evaluation of metagenome assemblies. Bioinformatics, 32(7), 1088–1090. https://doi.org/10.1093/bioinformatics/btv697

Moradigaravand, D., Palm, M., Farewell, A., Mustonen, V., Warringer, J., & Parts, L. (2018). Prediction of antibiotic resistance in Escherichia coli from large-scale pan-genome data. PLOS Computational Biology, 14(12), e1006258. https://doi.org/10.1371/journal.pcbi.1006258

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., Stepanauskas, R., McLean, J., Lasken, R., Clingenpeel, S. R., Woyke, T., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2013). Assembling genomes and mini-metagenomes from highly chimeric reads. In M. Deng, R. Jiang, F. Sun, & X. Zhang (Eds.), Research in Computational Molecular Biology (pp. 158–170). Springer Berlin Heidelberg.

Okonechnikov, K., Golosova, O., & Fursov, M. (2012). Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics, 28(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., & Stevens, R. (2013). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Research, 42(D1), D206–D214. https://doi.org/10.1093/nar/gkt1226

Owen, J. R., Noyes, N., Young, A. E., Prince, D. J., Blanchard, P. C., Lehenbauer, T. W., Aly, S. S., Davis, J. H., O'Rourke, S. M., Abdo, Z., Belk, K., Miller, M. R., Morley, P., & Van Eenennaam, A. L. (2017). Whole-genome sequencing and concordance between antimicrobial susceptibility genotypes and phenotypes of bacterial isolates associated with bovine respiratory disease. G3: Genes|Genomes|Genetics, 7(9), 3059–3071. https://doi.org/10.1534/g3.117.1137

Puchalski, A., Urban-Chmiel, R., Dec, M., Stęgierska, D., & Wernicki, A. (2016). The use of MALDI-TOF mass spectrometry for rapid identification of Mannheimia haemolytica. Journal of Veterinary Medical Science, 78(8), 1339–1342. https://doi.org/10.1292/jvms.16-0087

Samanta, I., & Bandyopadhyay, S. (2020). Pasteurella and Mannheimia. In Antimicrobial Resistance in Agriculture (pp. 263–273). Elsevier. https://doi.org/10.1016/B978-0-12-815770-1.00021-3

Seeger, J. T., King, M. E., Grotelueschen, D. M., Rogers, G. M., & Stokka, G. S. (2011). Effect of management, marketing, and certified health programs on the sale price of beef calves sold through a livestock video auction service from 1995 through 2009. Journal of the American Veterinary Medical Association, 239(4), 451–466. https://doi.org/10.2460/javma.239.4.451

Stanford, K., Zaheer, R., Klima, C., McAllister, T., Peters, D., Niu, Y. D., & Ralston, B. (2020). Antimicrobial resistance in members of the bacterial bovine respiratory disease complex isolated from lung tissue of cattle mortalities managed with or without the use of antimicrobials. Microorganisms, 8(2), 288. https://doi.org/10.3390/microorganisms8020288

Theurer, M. E., Johnson, M. D., Fox, T., McCarty, T. M., McCollum, R. M., Jones, T. M., & Alkire, D. O. (2021). Bovine respiratory disease during the mid-portion of the feeding period: Observations of frequency, timing, and population from the field. Applied Animal Science, 37(1), 52–58. https://doi.org/10.15232/aas.2020-02089

Torres-Blas, I., Fernández Aguilar, X., Cabezón, O., Aragon, V., & Migura-García, L. (2021). Antimicrobial resistance in Pasteurellaceae isolates from Pyrenean Chamois (Rupicapra pyrenaica) and domestic sheep in an alpine ecosystem. Animals, 11(6), 1686. https://doi.org/10.3390/ani11061686

Van Driessche, L., Bokma, J., Deprez, P., Haesebrouck, F., Boyen, F., & Pardon, B. (2019). Rapid identification of respiratory bacterial pathogens from bronchoalveolar lavage fluid in cattle by MALDI-TOF MS. Scientific Reports, 9(1), 18381. https://doi.org/10.1038/s41598-019-54599-9

Visick, J. E., Ichikawa, J. K., & Clarke, S. (1998). Mutations in the Escherichia coli surE gene increase isoaspartyl accumulation in a strain lacking the pcm repair methyltransferase but suppress stress-survival phenotypes. FEMS Microbiology Letters, 167(1), 19–25. https://doi.org/10.1111/j.1574-6968.1998.tb13202.x

Wood, T. L., & Wood, T. K. (2016). The HigB/HigA toxin/antitoxin system of Pseudomonas aeruginosa influences the virulence factors pyochelin, pyocyanin, and biofilm formation. MicrobiologyOpen, 5(3), 499–511. https://doi.org/10.1002/mbo3.346

CHAPTER 6. CONCLUSIONS AND FUTURE RESEARCH

6.1 Part I – Bacterial Community Analysis

6.1.1 The bacterial community in the reproductive tract of cattle has similar members to other mucosal-associated communities

Prior to the study discussed in Chapter 2 of this dissertation, there were no characterizations of the reproductive bacterial community in male livestock species. The bull preputial microbial community shares members with the bull urinary tract as well as the vaginal bacterial community of cows. The community seems to be made of bacterial members from soil, fecal, and vaginal communities, similarities that can also be found in other ruminant and non-ruminant species. While the impact of community composition on bull breeding was not studied, the differences between high and low diversity communities can be used in future research.

Following the examination of the bull prepuce from animals with satisfactory breeding soundness scores, a study on the composition and diversity of the seminal bacterial community of bulls with satisfactory and unsatisfactory scores was performed (Koziol et al., 2022). While the preputial community is made primarily of various soil, feces, and mucosal-associated bacteria, the seminal community of satisfactory bulls more closely resembled mucosal bacterial communities. The community of unsatisfactory bulls consisted of several bacterial members identified as opportunistic pathogens in cattle vaginal and respiratory tracts.

The identification of potential bacterial pathogens within the bull preputial and seminal communities leads to a few questions that can be used as the starting point of future research. First, is bacterial species richness, evenness, phylogenetic diversity, or a combination beneficial to prevent pathogenic colonization? Through the examination of the preputial community of healthy bulls, bacterial communities with low and high diversity with respect to the Shannon index were identified. The implication of these communities is still unknown and should be examined. In the vaginal community, cows considered to be “healthy” reproductively have low richness and phylogenetic diversity, but the individual members are well-established within the microbial community (Swartz et al., 2014).

Second, how does the seminal community affect the reproductive communities of the cow? The vaginal and uterine bacterial community composition have been studied, primarily

surrounding the effects of artificial insemination or pregnancy and parturition (Ault et al., 2019; Clemmons et al., 2017; Swartz et al., 2014). However, studies on the combination of seminal communities with different bacterial compositions and their effects on the vaginal community before and after breeding have not been performed and could shed some light on potential improvements for breeding programs.

Third, is there a relationship between the bull reproductive bacterial communities and cow reproductive diseases? Infections and diseases of the cow reproductive tract have been studied to find potential treatments and prophylactic solutions (Moore et al., 2021; Rosales & Ametaj, 2021). As with the first question, more needs to be done with the combination of seminal quality and the effects on the cow.

Last, following the previous two questions, are there long-lasting effects of the bull reproductive communities? The study(ies) would need to be longitudinal, examining different stages of reproduction of the cow (e.g., estrus, pregnancy, parturition, post-partum) and bull before and after breeding (i.e., directly bred with cows or collection of semen for artificial insemination), as well as following calves from the different breeding conditions. Reproductive disease in bulls and cows have been studied (Bellows et al., 2002; Brotman et al., 2014; Wagener et al., 2014), and some information is known about the types of risks associated with calving (Murray et al., 2016; Wenker et al., 2022), but the combined knowledge of the generational effects of sire and dam reproductive bacterial communities on calf health is lacking.

Generally, the study of bacterial communities in the reproductive tract of cattle has shown there are some members typically found in both the bull and the cow. However, as more information is obtained about the differences in bull semen, as well as uterine and vaginal bacterial communities, additional improvements can be made to breeding programs and cow-calf operations to reduce the incidence of disease, morbidity, and mortality.

6.1.2 Supplementation of soluble fiber in the piglet diet increases bacteria associated with short-chain fatty acid production

Weaning is a time of immense stress for animals. Not only does the diet change from liquid (milk) to solid feed, but often weaning is coupled with immediate transport to grow-finish operations. Supplementation of dextrin in the piglet diet pre-weaning showed the most promising changes in bacterial community members with differentially abundant bacteria associated with

short-chain fatty acid production compared to supplementation post-weaning. The abundance of SCFA-producing bacteria was coupled with an empirical increase in SCFA, mainly butyrate, in the pre-weaning supplementation groups.

While the results of the supplementing dextrin pre- and post-weaning were not substantial, the study highlighted some important impacts. First, pre-weaning dextrin supplementation led to more favorable bacterial taxa compared to post-weaning supplementation or no supplementation. Potential reasons for these outcomes is the selection of bacteria that can utilize dextrin before weaning. While research on the effects of pre-weaning dietary supplementation in pigs is limited, there have been studies examining the changes that occur in the gut microbial community during this time. A study by Frese et al. (2015) showed the drastic shift in diversity and functional potential of the microbial community of nursing vs weaned pigs. By introducing fiber early, the microbiota that developed during the pre-weaning period likely lingered as the animals were switched to solid feed.

However, it was found that supplementation of dextrin at any time resulted in better health outcomes and a favorable community when compared to pigs that received no dextrin at all. This study lends itself to future research in a few ways. The most limiting factor given the results from the study is how to incorporate the dextrin into the farrowing piglet diet. Dextrin supplementation pre-weaning was achieved by oral gavage, something not feasible in large, commercial operations. One potential method is to provide the supplemental dextrin in the water given to the animals. As dextrin is soluble in water, this would likely be the easiest to implement. In spite of the solubility, testing would need to be performed to determine the willingness of piglets to drink water with dissolved fiber compared to water alone.

The implementation of pre-weaning fiber leads to two other avenues of research. First, as there is not a practical way of controlling how much fiber an animal eats using *ad libitum* feeding, a study should be performed to determine the lowest concentration of fiber at which benefits can be consistently shown. Second, as mentioned above, this should be done alongside preferential testing to determine if there is a concentration at which the animals will not choose to eat or drink the supplementary feed, similar to recent studies on acceptability and preference of different copper supplementation (Villagómez-Estrada et al., 2020) and forages (Figueroa et al., 2020).

6.2 Part II – Bacterial Population Analysis

6.2.1 Characterization and analysis of antibiotic resistance in bovine respiratory disease pathogens is necessary to improve our ability of determining AR in the field

The three bovine respiratory disease (BRD) pathogens discussed in this dissertation, *Mannheimia haemolytica, Pasteurella multocida,* and *Histophilus somni*, are all considered opportunistic pathogens; they are typically only pathogenic during times of stress on the host immune system (Underwood et al., 2015). Though all three species are also members of the family *Pasteurellaceae,* the types of phenotypic and genotypic antibiotic resistances found differed. For this reason, it is imperative to continue collecting and reporting phenotypic resistance information alongside molecular and genomic characterizations of the individual pathogens.

The first BRD study discussed in this dissertation attempted to utilize the similarities between the three BRD pathogens to determine AR from the genome and was based on a previous examination by Owen et al. (2017). The authors analyzed the genomes of *M. haemolytica, P. multocida*, and *H. somni* using resistance databases and compared the results to phenotypic susceptibility tests of the same isolates to answer the question, "What is the applicability of resistance gene databases to three bacterial pathogens commonly implicated in BRD?"

Prediction of antibiotic resistance was improved with the machine learning method used. However, there were still questions as to why the resistance databases were not more predictive of resistance given their high accuracy in other pathogenic bacteria such as *Salmonella enterica* (Neuert et al., 2018). One potential answer is likely the most obvious; there are more studies on *Salmonella enterica* than the BRD pathogens.

In an effort to determine if studying a single species would allow more resolution of the antibiotic resistance prediction, *M. haemolytica* was used in the second BRD study discussed in this dissertation. Again, there was some improvement in prediction using the machine learning method, however, in this case it was determined that machine learning improved the ability to predict resistance in those antibiotics for which the resistance database prediction was low. This left more questions unanswered as four of the seven tested antibiotics had ≥ 90% accuracy of phenotypic prediction.

As a follow-up to looking at antibiotic resistances found within the genomes, both molecular and genetic serotyping should be performed. As discussed within the chapter on *M.*

haemolytica resistance prediction, there are two pathogenic serotypes and one commensal serotype of *M. haemolytica* commonly found within the respiratory tract of cattle. Genetic information conferring AR can easily passed between the commensal and pathogenic serotypes, making it worth having the knowledge of AR within the respiratory tract community. However, in an effort to elucidate the most likely causative agent in BRD cases, knowing the common serotype found within a farm, breed, or geographic region could lead to targeted vaccine development as described in a previous study by Andrés-Lasheras et al. (2019).

Ultimately, the study of bacterial ecology can help improve the lives of food production animals through the determination of beneficial community compositions and organisms, as well as the detrimental members. The utilization of multi-omics approaches (e.g., using host genomics, microbial genomics and transcriptomics, and metabolomics) to answer animal health questions can provide even deeper knowledge of the impacts of diet, environment, and stress on the host animal.

6.3 References

Andrés-Lasheras, S., Zaheer, R., Klima, C., Sanderson, H., Ortega Polo, R., Milani, M. R. M., Vertenten, G., & McAllister, T. A. (2019). Serotyping and antimicrobial resistance of Mannheimia haemolytica strains from European cattle with bovine respiratory disease. Research in Veterinary Science, 124, 10–12. https://doi.org/10.1016/j.rvsc.2018.12.021

Ault, T. B., Clemmons, B. A., Reese, S. T., Dantas, F. G., Franco, G. A., Smith, T. P. L., Edwards, J. L., Myer, P. R., & Pohler, K. G. (2019). Uterine and vaginal bacterial community diversity prior to artificial insemination between pregnant and nonpregnant postpartum cows1. Journal of Animal Science, 97(10), 4298–4304. https://doi.org/10.1093/jas/skz210

Bellows, D. S., Ott, S. L., & Bellows, R. A. (2002). Review: Cost of reproductive diseases and conditions in cattle. The Professional Animal Scientist, 18(1), 26–32. https://doi.org/10.15232/S1080-7446(15)31480-7

Brotman, R. M., Ravel, J., Bavoil, P. M., Gravitt, P. E., & Ghanem, K. G. (2014). Microbiome, sex hormones, and immune responses in the reproductive tract: Challenges for vaccine development against sexually transmitted infections. Vaccine, 32(14), 1543–1552. https://doi.org/10.1016/j.vaccine.2013.10.010

Clemmons, B. A., Reese, S. T., Dantas, F. G., Franco, G. A., Smith, T. P. L., Adeyosoye, O. I., Pohler, K. G., & Myer, P. R. (2017). Vaginal and uterine bacterial communities in postpartum lactating cows. Frontiers in Microbiology, 8, 1047. https://doi.org/10.3389/fmicb.2017.01047

Figueroa, J., del Río, K., Romero, F., Keim, J. P., & Gandarillas, M. (2020). Acceptability, preferences, and palatability of diets containing summer and winter brassica forage in growing pigs: a pilot study. Animals, 10(6), 1080. https://doi.org/10.3390/ani10061080

Frese, S. A., Parker, K., Calvert, C. C., & Mills, D. A. (2015). Diet shapes the gut microbiome of pigs during nursing and weaning. Microbiome, 3(1), 28. https://doi.org/10.1186/s40168-015-0091-8

Koziol, J. H., Sheets, T., Wickware, C. L., & Johnson, T. A. (2022). Composition and diversity of the seminal microbiota in bulls and its association with semen parameters. Theriogenology, 182, 17–25. https://doi.org/10.1016/j.theriogenology.2022.01.029

Moore, D. P., Cantón, G. J., & Louge Uriarte, E. L. (2021). Editorial: Infectious diseases affecting reproduction and the neonatal period in cattle. Frontiers in Veterinary Science, 8, 679007. https://doi.org/10.3389/fvets.2021.679007

Murray, C. F., Fick, L. J., Pajor, E. A., Barkema, H. W., Jelinski, M. D., & Windeyer, M. C. (2016). Calf management practices and associations with herd-level morbidity and mortality on beef cow-calf operations. Animal, 10(3), 468–477. https://doi.org/10.1017/S1751731115002062

Neuert, S., Nair, S., Day, M. R., Doumith, M., Ashton, P. M., Mellor, K. C., Jenkins, C., Hopkins, K. L., Woodford, N., de Pinna, E., Godbole, G., & Dallman, T. J. (2018). Prediction of phenotypic antimicrobial resistance profiles from whole genome sequences of non-typhoidal Salmonella enterica. Frontiers in Microbiology, 9, 592. https://doi.org/10.3389/fmicb.2018.00592

Owen, J. R., Noyes, N., Young, A. E., Prince, D. J., Blanchard, P. C., Lehenbauer, T. W., Aly, S. S., Davis, J. H., O'Rourke, S. M., Abdo, Z., Belk, K., Miller, M. R., Morley, P., & Van Eenennaam, A. L. (2017). Whole-genome sequencing and concordance between antimicrobial susceptibility genotypes and phenotypes of bacterial isolates associated with bovine respiratory disease. G3;#58; Genes|Genomes|Genetics, 7(9), 3059–3071. https://doi.org/10.1534/g3.117.1137

Rosales, E. B., & Ametaj, B. N. (2021). Reproductive tract infections in dairy cows: can probiotics curb down the incidence rate? Dairy, 2(1), 40–64. https://doi.org/10.3390/dairy2010004

Swartz, J. D., Lachman, M., Westveer, K., ONeill, T., Geary, T., Kott, R. W., Berardinelli, J. G., Hatfield, P. G., Thomson, J. M., Roberts, A., & Yeoman, C. J. (2014). Characterization of the vaginal microbiota of ewes and cows reveals a unique microbiota with low levels of Lactobacilli and near-neutral pH. Frontiers in Veterinary Science, 1. https://doi.org/10.3389/fvets.2014.00019

Underwood, W. J., Blauwiekel, R., Delano, M. L., Gillesby, R., Mischler, S. A., & Schoell, A. (2015). Biology and diseases of ruminants (sheep, goats, and cattle). In Laboratory Animal Medicine (pp. 623–694). Elsevier. https://doi.org/10.1016/B978-0-12-409527-4.00015-8

Villagómez-Estrada, S., Pérez, J. F., van Kuijk, S., Melo-Durán, D., Karimirad, R., & Solà-Oriol, D. (2020). Dietary preference of newly weaned pigs and nutrient interactions according to copper levels and sources with different solubility characteristics. Animals, 10(7), 1133. https://doi.org/10.3390/ani10071133

Wagener, K., Grunert, T., Prunner, I., Ehling-Schulz, M., & Drillich, M. (2014). Dynamics of uterine infections with Escherichia coli, Streptococcus uberis and Trueperella pyogenes in post-partum dairy cows and their association with clinical endometritis. The Veterinary Journal, 202(3), 527–532. https://doi.org/10.1016/j.tvjl.2014.08.023

Wenker, M. L., Verwer, C. M., Bokkers, E. A. M., te Beest, D. E., Gort, G., de Oliveira, D., Koets, A., Bruckmaier, R. M., Gross, J. J., & van Reenen, C. G. (2022). Effect of type of cow-calf contact on health, blood parameters, and performance of dairy cows and calves. Frontiers in Veterinary Science, 9, 855086. https://doi.org/10.3389/fvets.2022.855086

APPENDIX A. CHAPTER 2 CORRELEATION TABLES

All correlation tables in this appendix are organized as follows. First, each separate table shows the statistically significant (rho < -0.5 or rho > 0.5, q < 0.05) correlations of the relative abundance of one genus (OTU 1) with the relative abundance of other OTUs (OTU 2). Second, taxonomic classification columns listed are for OTU 2.

A.1. OTUs co-occurring with Bradyrhizobium.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0038	Otu0001	2.59E-09	1.02E-07	0.59	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0038	Otu0007	1.67E-07	6.58E-06	0.53	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0038	Otu0009	3.21E-08	1.26E-06	-0.56	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas
Otu0038	Otu0020	7.63E-07	3.00E-05	-0.51	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified

A.2. OTUs co-occurring with Fusobacterium.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0008	Otu0001	6.59E-13	2.59E-11	-0.69	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0008	Otu0007	2.22E-10	8.71E-09	-0.62	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0005	Otu0001	4.27E-10	1.68E-08	-0.62	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0005	Otu0007	1.60E-08	6.26E-07	-0.57	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0008	Otu0004	9.93E-07	3.90E-05	0.50	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0008	Otu0016	8.20E-07	3.22E-05	0.51	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0005	Otu0017	4.71E-07	1.85E-05	0.52	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia
Otu0008	Otu0051	4.65E-07	1.83E-05	0.52	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_unclassified
Otu0008	Otu0010	4.54E-07	1.78E-05	0.52	Mollicutes	Mycoplasmatales	Mycoplasmataceae	Mycoplasma
Otu0024	Otu0018	4.51E-07	1.77E-05	0.52	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0024	Otu0040	4.35E-07	1.71E-05	0.52	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0024	Otu0025	4.06E-07	1.59E-05	0.52	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified

Otu0008	Otu0023	2.68E-07	1.05E-05	0.53	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus
Otu0024	Otu0016	1.13E-07	4.42E-06	0.54	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0005	Otu0015	9.24E-08	3.63E-06	0.54	Actinobacteria	Actinomycetales	Actinomycetales_unclassified	Actinomycetales_unclassified
Otu0013	Otu0048	9.15E-08	3.59E-06	0.54	Clostridia	Clostridiales	Peptostreptococcaceae	Peptostreptococcus
Otu0008	Otu0019	7.63E-08	3.00E-06	0.55	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0005	Otu0003	4.65E-08	1.82E-06	0.55	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0024	Otu0035	4.47E-08	1.75E-06	0.55	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0024	Otu0027	4.31E-08	1.69E-06	0.55	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0008	Otu0017	4.09E-08	1.61E-06	0.56	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia
Otu0005	Otu0034	3.91E-08	1.54E-06	0.56	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0008	Otu0031	1.57E-08	6.15E-07	0.57	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0005	Otu0016	1.26E-08	4.93E-07	0.57	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0005	Otu0044	1.07E-08	4.21E-07	0.57	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus

Otu0024	Otu0054	1.05E-08	4.10E-07	0.58	Clostridia	Clostridiales	Peptostreptococcaceae	Filifactor
Otu0005	Otu0009	6.63E-09	2.60E-07	0.58	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas
Otu0008	Otu0009	5.41E-09	2.12E-07	0.58	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas
Otu0005	Otu0011	3.99E-09	1.57E-07	0.59	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0008	Otu0011	3.27E-09	1.28E-07	0.59	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0013	Otu0006	3.02E-09	1.18E-07	0.59	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0024	Otu0052	2.39E-09	9.39E-08	0.60	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0008	Otu0015	7.73E-10	3.04E-08	0.61	Actinobacteria	Actinomycetales	Actinomycetales_unclassified	Actinomycetales_unclassified
Otu0013	Otu0032	7.32E-10	2.87E-08	0.61	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0008	Otu0026	6.63E-10	2.60E-08	0.61	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter
Otu0013	Otu0034	6.11E-11	2.40E-09	0.64	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0024	Otu0070	5.59E-11	2.20E-09	0.64	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0024	Otu0004	2.33E-11	9.14E-10	0.65	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas

Otu0013	Otu0030	7.53E-12	2.96E-10	0.66	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0024	Otu0022	3.22E-13	1.26E-11	0.69	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotellaceae_unclassified
Otu0024	Otu0033	2.33E-13	9.16E-12	0.69	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0024	Otu0019	4.73E-14	2.11E-12	0.71	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0024	Otu0055	1.47E-14	8.14E-13	0.72	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0013	Otu0039	4.22E-15	2.62E-13	0.73	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0024	Otu0012	2.66E-15	1.69E-13	0.73	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0013	Otu0050	4.44E-16	2.89E-14	0.74	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0008	Otu0003	2.22E-16	1.45E-14	0.75	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0024	Otu0042	0	0	0.76	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified

A.3. OTUs co-occurring with Parvimonas.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0009	Otu0001	0	0	-0.78	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0009	Otu0007	0	0	-0.76	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0009	Otu0038	3.21E-08	1.26E-06	-0.56	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium
Otu0009	Otu0031	1.04E-07	4.08E-06	0.54	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0009	Otu0002	3.08E-08	1.21E-06	0.56	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	Histophilus
Otu0009	Otu0026	2.30E-08	9.03E-07	0.56	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter
Otu0009	Otu0005	6.63E-09	2.60E-07	0.58	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0009	Otu0051	5.96E-09	2.34E-07	0.58	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_unclassified
Otu0009	Otu0008	5.41E-09	2.12E-07	0.58	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0009	Otu0017	1.03E-09	4.06E-08	0.61	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia
Otu0009	Otu0003	8.89E-11	3.49E-09	0.63	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0009	Otu0016	2.71E-11	1.06E-09	0.65	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified

Otu0009	Otu0014	1.90E-11	7.46E-10	0.65	Bacilli	Bacillales	Bacillales_Incertae_Sedis_XI	Gemella
Otu0009	Otu0011	1.97E-12	7.72E-11	0.67	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0009	Otu0015	4.44E-16	2.89E-14	0.75	Actinobacteria	Actinomycetales	Actinomycetales_unclassified	Actinomycetales_unclassified
Otu0009	Otu0020	0	0	0.76	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified

A.4. OTUs co-occurring with Porphyromonas.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0004	Otu0001	7.50E-07	2.94E-05	-0.51	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0004	Otu0008	9.93E-07	3.90E-05	0.50	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0004	Otu0070	7.81E-07	3.07E-05	0.51	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0030	Otu0032	2.76E-07	1.08E-05	0.53	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0006	Otu0054	2.27E-07	8.91E-06	0.53	Clostridia	Clostridiales	Peptostreptococcaceae	Filifactor
Otu0006	Otu0048	2.04E-07	8.01E-06	0.53	Clostridia	Clostridiales	Peptostreptococcaceae	Peptostreptococcus
Otu0052	Otu0060	1.63E-07	6.40E-06	0.53	Actinobacteria	Actinomycetales	Actinomycetaceae	Actinomyces
Otu0004	Otu0021	6.36E-08	2.50E-06	0.55	Bacteroidia	Bacteroidales	Bacteroidales_unclassified	Bacteroidales_unclassified
Otu0052	Otu0012	2.31E-08	9.07E-07	0.56	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0039	Otu0030	2.14E-08	8.42E-07	0.57	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0030	Otu0039	2.14E-08	8.42E-07	0.57	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0004	Otu0018	2.07E-08	8.14E-07	0.57	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified

Otu0006	Otu0050	1.16E-08	4.57E-07	0.57	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0052	Otu0042	1.00E-08	3.94E-07	0.58	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0030	Otu0048	9.82E-09	3.86E-07	0.58	Clostridia	Clostridiales	Peptostreptococcaceae	Peptostreptococcus
Otu0004	Otu0023	7.65E-09	3.01E-07	0.58	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus
Otu0004	Otu0003	3.16E-09	1.24E-07	0.59	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0006	Otu0013	3.02E-09	1.18E-07	0.59	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0052	Otu0024	2.39E-09	9.39E-08	0.60	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0004	Otu0016	2.01E-09	7.89E-08	0.60	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0052	Otu0022	1.35E-09	5.29E-08	0.60	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotellaceae_unclassified
Otu0052	Otu0054	9.10E-10	3.58E-08	0.61	Clostridia	Clostridiales	Peptostreptococcaceae	Filifactor
Otu0052	Otu0025	8.24E-10	3.24E-08	0.61	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0004	Otu0027	3.14E-10	1.23E-08	0.62	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0039	Otu0040	3.07E-10	1.20E-08	0.62	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified

Otu0004	Otu0055	1.51E-10	5.92E-09	0.63	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0004	Otu0035	1.06E-10	4.17E-09	0.63	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0004	Otu0054	2.93E-11	1.15E-09	0.65	Clostridia	Clostridiales	Peptostreptococcaceae	Filifactor
Otu0004	Otu0024	2.33E-11	9.14E-10	0.65	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0039	Otu0048	1.19E-11	4.67E-10	0.66	Clostridia	Clostridiales	Peptostreptococcaceae	Peptostreptococcus
Otu0030	Otu0013	7.53E-12	2.96E-10	0.66	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0030	Otu0050	3.26E-12	1.28E-10	0.67	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0004	Otu0042	1.10E-12	4.32E-11	0.68	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0039	Otu0050	2.19E-13	8.59E-12	0.70	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0006	Otu0040	5.91E-14	2.55E-12	0.71	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0039	Otu0006	5.42E-14	2.38E-12	0.71	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0006	Otu0039	5.42E-14	2.38E-12	0.71	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0004	Otu0033	2.13E-14	1.11E-12	0.72	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified

Otu0052	Otu0040	1.47E-14	8.14E-13	0.72	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonadaceae_unclassified
Otu0039	Otu0013	4.22E-15	2.62E-13	0.73	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0004	Otu0012	8.88E-16	5.75E-14	0.74	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0004	Otu0019	0	0	0.79	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides

A.5. OTUs co-occurring with Unclassified Ruminococcaceae.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0020	Otu0001	1.09E-13	4.32E-12	-0.70	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0020	Otu0007	7.86E-13	3.09E-11	-0.68	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0016	Otu0001	1.22E-11	4.77E-10	-0.66	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0016	Otu0007	4.59E-09	1.80E-07	-0.59	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0020	Otu0038	7.63E-07	3.00E-05	-0.51	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae	Bradyrhizobium
Otu0016	Otu0008	8.20E-07	3.22E-05	0.51	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0016	Otu0003	6.61E-07	2.59E-05	0.51	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0016	Otu0010	2.38E-07	9.36E-06	0.53	Mollicutes	Mycoplasmatales	Mycoplasmataceae	Mycoplasma
Otu0020	Otu0044	2.10E-07	8.26E-06	0.53	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus
Otu0016	Otu0024	1.13E-07	4.42E-06	0.54	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0016	Otu0011	1.02E-07	4.01E-06	0.54	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0016	Otu0018	9.52E-08	3.74E-06	0.54	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified

Otu0020	Otu0016	4.79E-08	1.88E-06	0.55	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0016	Otu0020	4.79E-08	1.88E-06	0.55	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0016	Otu0021	4.57E-08	1.79E-06	0.55	Bacteroidia	Bacteroidales	Bacteroidales_unclassified	Bacteroidales_unclassified
Otu0016	Otu0026	3.60E-08	1.41E-06	0.56	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter
Otu0020	Otu0026	1.34E-08	5.28E-07	0.57	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter
Otu0016	Otu0005	1.26E-08	4.93E-07	0.57	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0016	Otu0044	1.17E-08	4.58E-07	0.57	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus
Otu0016	Otu0019	7.25E-09	2.85E-07	0.58	Bacteroidia	Bacteroidales	Bacteroidaceae	Bacteroides
Otu0020	Otu0002	6.42E-09	2.52E-07	0.58	Gammaproteobacteria	Pasteurellales	Pasteurellaceae	Histophilus
Otu0016	Otu0070	5.98E-09	2.35E-07	0.58	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0016	Otu0023	3.19E-09	1.25E-07	0.59	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus
Otu0016	Otu0004	2.01E-09	7.89E-08	0.60	Bacteroidia	Bacteroidales	Porphyromonadaceae	Porphyromonas
Otu0016	Otu0033	1.15E-09	4.52E-08	0.60	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified

Otu0016	Otu0042	3.68E-10	1.44E-08	0.62	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0020	Otu0003	3.54E-10	1.39E-08	0.62	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0020	Otu0014	2.76E-10	1.08E-08	0.62	Bacilli	Bacillales	Bacillales_Incertae_Sedis_XI	Gemella
Otu0020	Otu0017	1.31E-10	5.12E-09	0.63	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia
Otu0016	Otu0055	4.58E-11	1.80E-09	0.64	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0016	Otu0009	2.71E-11	1.06E-09	0.65	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas
Otu0020	Otu0051	2.61E-11	1.03E-09	0.65	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_unclassified
Otu0016	Otu0111	1.53E-11	6.02E-10	0.65	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0020	Otu0011	1.01E-12	3.96E-11	0.68	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0020	Otu0015	0	0	0.76	Actinobacteria	Actinomycetales	Actinomycetales_unclassified	Actinomycetales_unclassified
Otu0020	Otu0009	0	0	0.76	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas

A.6. OTUs co-occurring with Streptobacillus.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0044	Otu0001	3.64E-08	1.43E-06	-0.56	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0044	Otu0007	2.33E-07	9.17E-06	-0.53	Bacilli	Bacillales	Bacillaceae_1	Bacillus
Otu0044	Otu0018	7.41E-07	2.91E-05	0.51	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0044	Otu0051	5.63E-07	2.21E-05	0.51	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_unclassified
Otu0044	Otu0020	2.10E-07	8.26E-06	0.53	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0044	Otu0026	8.26E-08	3.24E-06	0.55	Epsilonproteobacteria	Campylobacterales	Campylobacteraceae	Campylobacter
Otu0044	Otu0111	6.36E-08	2.50E-06	0.55	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0044	Otu0016	1.17E-08	4.58E-07	0.57	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0044	Otu0005	1.07E-08	4.21E-07	0.57	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0044	Otu0003	4.92E-09	1.93E-07	0.59	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0044	Otu0010	3.74E-09	1.47E-07	0.59	Mollicutes	Mycoplasmatales	Mycoplasmataceae	Mycoplasma
Otu0044	Otu0011	1.00E-10	3.93E-09	0.63	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu0044	Otu0017	1.29E-11	5.06E-10	0.66	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia

A.7. OTUs co-occurring with Histophilus.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu00 02	Otu00 01	4.20E-08	1.65E-06	-0.56	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu00 02	Otu00 11	8.01E-08	3.14E-06	0.55	Clostridia	Clostridiales	Lachnospiraceae	Lachnospiraceae_unclassified
Otu00 02	Otu00 17	6.74E-08	2.65E-06	0.55	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichia
Otu00 02	Otu00 09	3.08E-08	1.21E-06	0.56	Clostridia	Clostridiales	Clostridiales_Incertae_Sedis_XI	Parvimonas
Otu00 02	Otu00 43	2.98E-08	1.17E-06	0.56	Mollicutes	Mycoplasmatales	Mycoplasmataceae	Ureaplasma
Otu00 02	Otu00 20	6.42E-09	2.52E-07	0.58	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu00 02	Otu00 51	1.75E-09	6.86E-08	0.60	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Erysipelotrichaceae_unclassified
Otu00 02	Otu00 15	6.92E-11	2.72E-09	0.64	Actinobacteria	Actinomycetales	Actinomycetales_unclassified	Actinomycetales_unclassified
Otu00 02	Otu00 29	3.30E-11	1.30E-09	0.65	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified

A.8. OTUs co-occurring with Mycoplasma.

OTU 1	OTU 2	pval	qval	rho	Class	Family	Order	Genus
Otu0010	Otu0001	2.30E-07	9.05E-06	-0.53	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Escherichia
Otu0010	Otu0031	9.92E-07	3.89E-05	0.50	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0010	Otu0018	5.57E-07	2.19E-05	0.51	Clostridia	Clostridiales	Clostridiales_unclassified	Clostridiales_unclassified
Otu0010	Otu0008	4.54E-07	1.78E-05	0.52	Fusobacteriia	Fusobacteriales	Fusobacteriaceae	Fusobacterium
Otu0078	Otu0054	2.72E-07	1.07E-05	0.53	Clostridia	Clostridiales	Peptostreptococcaceae	Filifactor
Otu0010	Otu0016	2.38E-07	9.36E-06	0.53	Clostridia	Clostridiales	Ruminococcaceae	Ruminococcaceae_unclassified
Otu0078	Otu0033	1.20E-07	4.73E-06	0.54	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0010	Otu0003	6.36E-08	2.50E-06	0.55	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Leptotrichiaceae_unclassified
Otu0078	Otu0042	5.12E-08	2.01E-06	0.55	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified
Otu0010	Otu0044	3.74E-09	1.47E-07	0.59	Fusobacteriia	Fusobacteriales	Leptotrichiaceae	Streptobacillus

APPENDIX B. CHAPTER 4 SUPPLEMENTARY TABLES AND FIGURES

B.1. Assembly statistics for Purdue ADDL isolates

Species	Isolate	Contigs	Coverage	Total Length	N50	Max Contig Length
HS	A197597	17	94.61	2084808	237695	556989
HS	A198251	32	103.81	2287988	201957	323604
HS	A198252	31	103.79	2287280	163145	477970
HS	A198296	22	97.97	2158031	290280	445746
HS	B6291	18	97.81	2153613	187280	590911
MH	A194411	63	99.09	2689165	108982	339662
MH	A194414	68	99.11	2689066	108982	230838
MH	A194420	79	99.19	2692388	108963	339547
MH	A194424	68	99.17	2689916	102105	262623
MH	A196038	62	96.96	2625639	102146	339528
MH	A198447	75	96.80	2627581	102098	339547
PM	A193010	25	105.07	2318066	216932	537536
PM	A196291	36	111.25	2453934	199122	356656
PM	A196714	66	211.43	4666579	285797	754889
PM	A197897	24	107.60	2373074	206593	423786
PM	A198602	22	105.49	2323030	194751	423790
PM	A198640	15	100.78	2218602	534196	608605
PM	A199107	39	109.54	2415739	198905	318006

B.2. Minimum inhibitory concentrations of antibiotics for Purdue PA isolates. All concentrations in mg/mL

Isolate	Species	Ampicillin	Ceftiofur	Clindamycin	Danofloxacin	Enrofloxacin	Florfenicol	Gentamicin	Neomycin	Penicillin	Spectinomycin	Sulfadimethoxine	Tiamulin	Tilmicosin	Trimeth/sulfa	Tulathromycin	Tylosin tartrate
A193010	*P. multocida*	0.25	0.25	8	0.12	0.12	0.25	2	8	0.12	8	256	8	2	2	8	8
A194411	*M. haemolytica*	0.5	0.25	16	1	2	8	16	32	1	64	256	8	16	2	64	32
A194414	*M. haemolytica*	0.5	0.25	16	1	2	8	16	32	1	64	256	8	16	2	64	32
A194420	*M. haemolytica*	16	0.25	16	1	2	8	16	32	8	64	256	16	16	2	64	32
A194424	*M. haemolytica*	0.5	0.25	16	1	2	8	16	32	4	64	256	8	16	2	64	32
A196038	*M. haemolytica*	0.25	0.25	8	0.12	0.12	0.5	2	4	0.12	32	256	8	4	2	8	32
A196291	*P. multocida*	0.25	0.25	8	0.12	0.12	0.25	2	32	0.12	64	256	16	2	2	8	16
A196714	*P. multocida*	16	0.5	16	1	2	8	2	32	8	64	256	32	16	2	64	32
A197597	*H. somni*	0.25	0.25	1	0.12	0.12	0.25	4	16	0.12	16	256	2	4	2	8	4
A197897	*P. multocida*	0.25	0.25	8	0.12	0.25	0.25	4	8	0.12	8	256	8	2	2	8	8
A198251	*H. somni*	0.25	0.25	2	1	1	1	4	32	0.12	64	256	1	8	2	16	2
A198252	*H. somni*	0.25	0.25	16	1	1	2	4	32	0.12	64	256	1	16	2	32	4
A198296	*H. somni*	0.25	0.25	1	0.12	0.12	0.5	8	32	0.12	16	256	2	4	2	8	4
A198447	*M. haemolytica*	0.25	0.25	8	0.12	0.12	0.5	2	4	0.25	16	256	8	4	2	8	32
A198602	*P. multocida*	0.25	0.25	0.25	0.12	0.12	0.25	2	32	0.12	8	256	1	2	2	8	4
A198640	*P. multocida*	0.25	0.25	16	0.12	0.12	0.5	2	4	0.12	16	256	16	8	2	8	32
A199107	*P. multocida*	0.25	0.25	16	0.12	0.12	0.25	4	32	0.12	64	256	16	4	2	8	16
B6291	*H. somni*	0.25	0.25	1	0.12	0.12	0.25	16	32	0.12	32	256	2	4	2	8	4

B.3. Sequence information for Rule alignments. Length of region refers to length of filtered alignment (Reference) or rule match (Region).

		Length of region	% Similarity to alignment	% Isolates with matches
Danofloxacin, 0.5mg				
Rule 1 - Presence	Region 1	61	97-100%	100%
	Reference	61		
Rule 2 - Absence	Region 1	62	82-100%	100.00%
	Reference	62		
Tilmicosin, 16 mg				
Rule 1 - Absence	Region 1	71	95-100%	96%
	Region 2	34	91-100%	96%
	Region 3	40	95-100%	96%
	Reference	744		
Tulathromycin, 32 mg				
Rule 1 - Presence	Region 1	50	98-100%	100%
	Reference	50		
Rule 2 - Absence	Region 1	60	100%	42%
	Reference	60		
Ampicillin, 16 mg				
Rule 1 - Presence	Region 1	47	100%	97%
	Reference	47		
Rule 2 - Presence	Region 1	67	100%	95%
	Region 2	61	96-100%	100%
	Reference	128		
Clindamycin, 8mg				
Rule 1 - Absence	Region 1	61	100%	100%
	Region 2	41	100%	100%
	Reference	705		
Rule 2 - Presence	Region 1	32	87-100%	100%
	Reference	32		
Penicillin, 1 mg				
Rule 1 - Presence	Region 1	62	55-100%	100%
	Reference	62		
Tiamulin, 8 mg				
Rule 1 - Absence	Region 1	7369	10-96%	100%
	Region 2	5148	27-60%	100%
	Reference	Several regions		

B.4. Antibiotic resistance genes per isolate. '+': gene present in isolate; '-': gene absent from isolate. MH: *Mannheimia haemolytica*; PM*: Pasteurella multocida*; HS: *Histophilus somni;* PA: Purdue ADDL isolates; OA: Owen isolates

			MLS				Phenicol		Tetracycline		Aminoglycosides					Trimpethoprim	Sulfonamide	Beta-lactam		
Isolate	Species	Source	*Erm(42)*	*ErmF*	*mphE*	*msrE*	*floR*	*Ccol_ACT_CHL*	*tet(H)*	*tet(D)*	*aadB*	*APH(6)-Id (strB)*	*APH(3")-Ib (strA)*	*aadA25*	*aadB*	*dfrA14*	*sul2*	*ROB-1*	*OXA-2*	*ROB-2*
A193010	*PM*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A194411	*MH*	PA	+	-	+	+	+	-	+	-	+	+	+	+	+	-	+	-	+	-
A194414	*MH*	PA	+	-	+	+	+	-	+	-	+	+	+	+	+	-	+	-	+	-
A194420	*MH*	PA	+	-	+	+	+	-	+	-	+	+	+	+	+	-	+	+	+	-
A194424	*MH*	PA	+	-	+	+	+	-	+	-	+	+	+	+	+	-	+	-	+	-
A196038	*MH*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A196291	*PM*	PA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	-	-
A196714	*PM*	PA	-	+	+	+	+	-	-	+	+	+	-	-	-	+	+	-	-	+
A197597	*HS*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A197897	*PM*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A198251	*HS*	PA	+	-	+	+	+	-	+	-	+	+	+	+	-	-	+	-	+	-
A198252	*HS*	PA	+	-	+	+	+	-	+	-	+	+	+	+	-	-	+	-	+	-
A198296	*HS*	PA	-	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
A198447	*MH*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A198602	*PM*	PA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	-	-
A198640	*PM*	PA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
A199107	*PM*	PA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	-	-
B6291	*HS*	PA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM192945v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	+	-	-
ASM192946v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-

B.4. Continued			MLS				Phenicol		Tetracycline		Aminoglycoside					Trimethoprim	Sulfonamide	Beta-lactam		
Isolate	Species	Source	*Erm(42)*	*ErmF*	*mphE*	*msrE*	*floR*	*Ccol_ACT_CHL*	*tet(H)*	*tet(D)*	*aadB*	*APH(6)-Id (strB)*	*APH(3'')-Ib (strA)*	*aadA25*	*aadB*	*dfrA14*	*sul2*	*ROB-1*	*OXA-2*	*ROB-2*
ASM192951v1	*MH*	OA	-	-	-	-	-	+	+	-	+	+	+	-	-	-	-	-	-	-
ASM192952v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192956v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
ASM192959v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192964v1	*MH*	OA	-	-	-	-	-	+	+	-	+	+	+	-	-	-	-	-	-	-
ASM192965v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192966v1	*MH*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM192970v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	+	-	-	-	-	-	-	-
ASM192971v1	*MH*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192973v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192976v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM192977v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	+	-	-
ASM192985v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM192988v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	-	+	-	-
ASM192989v1	*HS*	OA	+	+	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM192990v1	*PM*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM192996v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	-	-	-	-	-	+	-	-
ASM192998v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM193002v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	-	+	-	-
ASM193010v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	-	+	-	-
ASM193019v1	*MH*	OA	-	-	-	-	-	+	+	-	+	-	-	-	-	-	-	+	-	-
ASM193028v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM193030v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-

B.4. Continued			MLS				Phenicol		Tetracycline		Aminoglycoside					Trimethoprim	Sulfonamide	Beta-lactam		
Isolate	Species	Source	*Erm(42)*	*ErmF*	*mphE*	*msrE*	*floR*	*Ccol_ACT_CHL*	*tet(H)*	*tet(D)*	*aadB*	*APH(6)-Id (strB)*	*APH(3")-Ib (strA)*	*aadA25*	*aadB*	*dfrA14*	*sul2*	*ROB-1*	*OXA-2*	*ROB-2*
ASM193032v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	+	-	-
ASM193036v1	*MH*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM193038v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
ASM193040v1	*PM*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM193044v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM193046v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	+	-	-	-	-	-	-	-
ASM193047v1	*MH*	OA	+	-	-	-	-	-	+	-	+	-	+	-	-	-	-	-	-	-
ASM193050v1	*MH*	OA	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-
ASM193058v1	*MH*	OA	-	-	-	-	-	+	+	-	+	-	-	-	-	-	-	-	-	-
ASM193060v1	*PM*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM193061v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	+	-	-
ASM193062v1	*MH*	OA	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-
ASM193068v1	*MH*	OA	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-
ASM193069v1	*MH*	OA	-	-	-	-	-	-	+	-	+	-	+	-	-	-	-	+	-	-
ASM193070v1	*PM*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM193073v1	*MH*	OA	+	-	-	-	+	-	+	-	+	+	+	-	-	-	+	-	-	-
ASM193076v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM193077v1	*HS*	OA	+	-	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM193078v1	*HS*	OA	+	+	-	-	-	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM193082v1	*HS*	OA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ASM193084v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	+	-	-
ASM193087v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	+	-	-	-	-	-	-	-
ASM193089v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-

B.4. Continued			MLS				Phenicol		Tetracycline		Aminoglycoside					Trimethoprim	Sulfonamide	Beta-lactam		
Isolate	Species	Source	*Erm(42)*	*ErmF*	*mphE*	*msrE*	*floR*	*Ccol_ACT_CHL*	*tet(H)*	*tet(D)*	*aadB*	*APH(6)-Id (strB)*	*APH(3'')-Ib (strA)*	*aadA25*	*aadB*	*dfrA14*	*sul2*	*ROB-1*	*OXA-2*	*ROB-2*
ASM193092v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM193093v1	*MH*	OA	-	-	-	-	-	-	+	-	+	-	-	-	-	-	-	+	-	-
ASM193096v1	*PM*	OA	+	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	-	-
ASM193097v1	*HS*	OA	-	-	-	-	-	-	+	+	+	+	+	-	-	-	+	-	-	-
ASM193102v1	*HS*	OA	+	+	-	-	-	-	+	-	+	+	-	-	-	+	-	-	-	-
ASM193103v1	*HS*	OA	+	+	-	-	-	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM193104v1	*HS*	OA	+	+	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM193108v1	*HS*	OA	+	+	-	-	+	-	+	-	+	-	-	-	-	+	-	-	-	-
ASM193109v1	*HS*	OA	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	-
ASM193123v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	+	-	-	-	-	-	-	-
ASM193126v1	*PM*	OA	+	-	-	-	-	-	+	-	+	+	-	-	-	-	-	-	-	-
ASM193134v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	-	+	-	-
ASM193136v1	*MH*	OA	-	-	-	-	-	-	+	-	+	+	+	-	-	-	+	+	-	-
ASM193137v1	*HS*	OA	-	-	-	-	-	-	-	+	+	+	-	-	-	-	-	-	-	-
ASM193141v1	*HS*	OA	-	-	-	-	-	-	-	+	+	+	+	-	-	-	-	-	-	-
ASM193144v1	*HS*	OA	+	+	-	-	-	-	+	-	+	-	-	-	-	+	-	-	-	-

B.5. Annotations for Rule rulesets. Functional annotation of region refers to BLAST annotation of filtered alignment (BLAST reference) or RAST annotation of isolate match (Region)

Danofloxacin, 0.5mg		Functional annotation of region	Interesting Surrounding genes	
Rule 1 – Presence	Region 1	mobile element protein	Integrase, APH genes, sul2, floR, type 4 pilus/sec system	Several found as only CDS on contig
	BLAST Reference	IS30-like element ISApl1 family transposase	Integrase, APH genes, aminopenicillin acyl-transferase	*Pasteurellaceae* other than MH, PM, HS
Rule 2 – Absence	Region 1	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	Stress proteins and ABC transporters	Hits with e-values 10^-4 (MH) to 10^-30 (HS)
	BLAST Reference	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA)	LPS trasnporters, exoribonuclease gene, NADH reductase	All references were *PM*
Tilmicosin, 16 mg				
Rule 1 – Absence	Region 1	23S	If any surrounding genes, other rRNA (16S and 5S)	E-values around 10^-33
	Region 2	23S		E-values aroud 10^-13
	Region 3	23S		E-values around 10^17
	BLAST Reference	23S	Surrounded by tRNAs	References are mostly other *Pasteuerellaceae* (*Glaesserella, Avibacterium, Haemophilus*)
Tulathromycin, 32 mg				
Rule 1 – Presence	Region 1	region between tetR and tetH	integrase, 184ransposase, metal resistance, other ARG	E-values around 10^-22, no hits to CDS
	BLAST Reference	region between tetR and tetH	ISL3, metal resistance	Matches to many ICE assemblies
Rule 2 – Absence	Region 1	23S rRNA specific to *Pasteurella*	23S rRNA	All isolates with hits were *PM*
	BLAST Reference	23S rRNA specific to *Pasteurella*	23S rRNA	All references were *PM*

B.5. Continued

Ampicillin, 16 mg				
Rule 1 - Presence	Region 1	Histidinol transaminase	Cell division machinery, specific sugar metabolism, aa synth genes	E-values of 10^-21
	BLAST Reference	Histidinol transaminase	aa synth genes, metalloprotease	Most references *MH,* some *Bibersteinia*, and *Pasterurellaceae* bacterium
Rule 2 - Presence	Region 1	hypothetical protein, or region before/after/between other CDS	integrases, mobile element/mobilization proteins	One MH had no hit for region 1, but had a hit for region 2
	Region 2	mobile element protein		
	BLAST Reference	IS30-like element ISApl1 family transposase	sul2, tetB, tetR, hth transcriptional regulator	Most references *MH,* some *Bibersteinia*, and *Pasterurellaceae* bacterium
Clindamycin, 4mg				
Rule 1 - Absence	Region 1	23S rRNA	If any surrounding genes, other rRNA (16S and 5S)	Regions matched 23S rRNA on same contig, often truncated
	Region 2			
	BLAST Reference	23S rRNA	other rRNA genes and replication genes	References are mostly other *Pasteuerellaceae* (*Glaesserella, Avibacterium, Haemophilus*) with several *HS*
Rule 2 - Presence	Region 1	tRNA-Pro-TGG	Other tRNA, aa synthesis, and protein metab genes	Matched 2 copies of tRNA-Pro with anticodon TGG for each genome
	BLAST Reference	tRNA-Pro-TGG	Other tRNAs and rRNA genes	References are mostly other *Pasteuerellaceae* (*Glaesserella, Avibacterium, Haemophilus*)
Penicillin, 1 mg				
Rule 1 - Presence	Region 1	Hisitidnol dehydrogenase	membrane proteins, purine biosynthesis	PM isolates matched correct gene but had high e-values
	BLAST Reference	Histidinol dehyrogenase	membrane proteins	Majority of references were MH

B.5. Continued

Tiamulin, 32 mg				
Rule 1 - Presence	Region 1	CRISPR array	LPS synth, sugar metabolism (PTS, LacI), RNA pol genes, RNA helicase	Repeat region only found in *PM* isolates
	BLAST Reference	CRISPR repeat region	LPS synth, sugar metabolism (PTS, LacI), RNA pol genes, RNA helicase	All references PM

APPENDIX C. CHAPTER 5 SUPPLEMENTARY TABLES

C.1. Assembly statistics

Isolate ID	# contigs	Max contig (bp)	Total length (bp)	N50
MH001	88	339547	2681281	95744
MH002	80	230826	2585295	87508
MH003	101	208386	2620333	78085
MH004	117	154831	2552581	69710
MH005	77	339498	2612630	111021
MH006	113	339547	2706083	85681
MH007	107	184143	2687938	76013
MH008	94	339548	2688989	87630
MH009	81	339547	2651437	94189
MH010	112	201624	2685788	57726
MH011	489	62800	2595262	8056
MH012	37	439203	2435732	180083
MH013	89	333432	2680800	87630
MH014	76	339335	2587889	102104
MH015	89	339548	2608987	89064
MH016	105	333410	2676824	74011
MH017	101	339547	2687866	76225
MH018	105	150048	2524369	62189
MH019	59	339460	2568758	126573
MH020	76	339309	2589273	102104
MH021	79	339535	2582633	97035
MH022	79	333382	2643712	97035
MH023	99	230823	2694353	92346
MH024	97	339334	2622499	102105
MH025	98	339529	2594939	95378
MH026	108	149813	2519133	64501
MH027	93	339516	2629826	95378
MH028	89	339546	2628628	84801
MH029	88	333417	2661999	93126
MH030	90	345738	2619733	95378
MH031	81	339467	2677174	97035
MH032	81	339534	2652343	94188
MH033	107	155103	2526032	69701

C.1. continued

Isolate ID	# contigs	Max contig (bp)	Total length (bp)	N50
MH034	102	339511	2626503	82739
MH035	99	331010	2675934	87635
MH036	95	339535	2614743	95378
MH037	88	339493	2619963	95378
MH038	89	339496	2628160	97035
MH039	105	330885	2608996	87640
MH040	99	216941	2627213	95378
MH042	86	333465	2685515	94189
MH043	97	339511	2625181	97035
MH044	97	339548	2627964	78313
MH045	64	333390	2569523	126617
MH046	85	339433	2569630	97035
MH047	97	333476	2622612	78142
MH048	80	230856	2624990	102104
MH049	104	216941	2607304	82958
MH050	102	339505	2626699	95378
MH052	117	330944	2711544	87715
MH053	89	330992	2657884	89205
MH054	95	330957	2687240	95744
MH055	92	333476	2688315	97035
MH056	95	330903	2683535	87635
MH057	90	330992	2657097	87635
MH058	90	230841	2657509	89205
MH059	92	333467	2683594	89205
MH060	82	333324	2621609	97035
MH061	93	330974	2684370	87715
MH062	95	330980	2689908	95745
MH063	88	208386	2656731	89205
MH064	94	330949	2652165	89214
MH065	94	330908	2688869	89214
MH066	90	208386	2689725	89223
MH067	76	333312	2643484	102095
MH068	90	230932	2647843	102105
MH069	106	330810	2650647	89184
MH070	88	230842	2648691	102129
MH071	101	330956	2647583	97035

C.1. continued

Isolate ID	# contigs	Max contig (bp)	Total length (bp)	N50
MH072	108	143862	2550508	66206
MH073	110	208386	2687760	87635
MH074	85	330966	2681643	89187
MH075	100	333353	2624743	87640
MH076	105	155402	2676786	68332
MH077	83	333410	2583633	89205
MH078	102	149132	2517053	78321
MH079	99	330950	2622216	89250
MH080	106	333470	2624942	82966
MH081	85	330992	2582616	87722
MH082	43	346448	2352287	175491
MH083	102	149001	2517489	77688
MH084	123	443660	4939555	86453
MH085	90	210870	2625518	89241
MH086	102	330938	2624683	87640
MH087	103	185695	2516557	69643
MH088	109	333476	2660837	89250
MH2125	107	152771	2659536	72146
MH2165	115	168440	2659886	55668
MH2190	129	168440	2671976	60237
MH2297	389	100950	2907601	32362
MH2428	107	155255	2663675	61400
MH2436	114	155287	2605887	74883
MH2512	140	149596	2582071	61082
MH2543	104	155257	2594941	93925
MH2578	116	168439	2606708	80767
MH2597	105	168439	2595575	80767
MH2612	136	168440	2656249	53418
MH2633	147	140982	2608186	60589
MH2683	125	155288	2658889	61400
MH2700	135	168446	2657234	60237
MH2769	127	140982	2573893	61006
MH2887	102	153194	2658824	73859
MH2930	125	168440	2659136	53418
MH2969	144	183847	2676978	56429
MH3022	131	168440	2657638	60237

C.1. continued

Isolate ID	# contigs	Max contig (bp)	Total length (bp)	N50
MH3045	100	155254	2600854	91992
MH3369	132	168441	2657754	59910
MH23209	139	149585	2598318	56752
MH32754	142	141405	2589936	56748
MH232020	120	168440	2657117	73858
MH327511	110	150867	2656656	61400
MH327518	143	141155	2582644	50296

C.2. Minimum inhibitory concentrations of antibiotics for ADDL, KSU, and TAMU isolates used in machine learning model creation. All concentrations in mg/mL.

AMP = ampicillin; CEFT=ceftiofur; CLIND = clindamycin; DANO = danofloxacin; ENRO = enrofloxacin; FLOR = florfenicol; GAMY = gamythromycin; GENT = gentamicin; NEO = neomycin; OTET = oxytetracycline; PEN = penicillin; SPEC = spectinomycin; SULFA = sulfadimethoxine; TET = tetracycline; TIA = tiamulin; TILD = tildipirosin; TILM = tilmicosin; TMS = trimethoprim/sulfamethoxazole; TULA = tulathromycin; TYL = tylosin tartrate

Isolate	AMP	CEFT	CLIND	DANO	ENRO	FLOR	GAMY	GENT	NEO	OTET	PEN	SPEC	SULFA	TET	TIA	TILD	TILM	TMS	TULA	TYL
A194411	0.5	0.25	16	1	2	8		16	32		1	64	256		8		16	2	64	32
A194414	0.5	0.25	16	1	2	8		16	32		1	64	256		8		16	2	64	32
A194420	16	0.25	16	1	2	8		16	32		8	64	256		16		16	2	64	32
A194424	0.5	0.25	16	1	2	8		16	32		4	64	256		8		16	2	64	32
A196038	0.25	0.25	8	0.12	0.12	0.5		2	4		0.12	32	256		8		4	2	8	32
A198447	0.25	0.25	8	0.12	0.12	0.5		2	4		0.25	16	256		8		4	2	8	32
MH001	0.25	0.25	16	1	2	8		16	32	8	8	0.5	64	4	256		4	8		64
MH002	0.25	0.25	16	0.12	0.12	1		2	8	1	1	0.25	32	0.5	256		0.5	16		4
MH004	0.25	0.25	8	0.12	0.12	0.5		2	4	0.5	0.5	0.25	32	0.5	256		0.5	16		8
MH006	0.25	0.25	16	1	2	8		16	32	8	8	0.5	64	8	256		8	8		64
MH008	0.25	0.25	16	1	2	8		16	32	8	8	0.5	64	8	256		8	8		64
MH009	0.25	0.25	8	1	2	1		16	4	8	8	0.5	64	8	256		8	8		32
MH010	16	0.5	16	1	2	8		16	32	8	8	8	64	8	256		8	8		64
MH011	0.25	0.25	8	0.12	0.12	0.5		2	4	0.5	0.5	0.12	32	0.5	256		0.5	8		4
MH012	0.25	2	4	0.5	0.5	8		1	16	8	8	0.12	8	8	256		8	16		4
MH013	0.25	0.25	16	1	2	8		1	32	8	8	0.25	32	8	256		8	8		64
MH015	0.25	0.25	8	0.12	0.12	0.5				0.5	0.5	0.25	32	0.5	256		0.5	16		4
MH016	0.25	0.25	16	1	2	2				8	8	0.25	64	8	256		8	8		64
MH018	0.25	0.25	8	0.12	0.12	0.5				0.5	0.5	0.25	32	0.5	256		0.5	16		8
MH019	0.25	0.25	16	0.12	0.12	0.5				8	8	0.25	32	8	256		8	16		8

C.2. continued

Isolate	AMP	CEFT	CLIND	DANO	ENRO	FLOR	GAMY	GENT	NEO	OTET	PEN	SPEC	SULFA	TET	TIA	TILD	TILM	TMS	TULA	TYL
MH021	0.25	0.25	8	0.12	0.12	1				0.5	0.5	0.12	16	1	256		1	16		4
MH023	16	0.25	16	1	2	8				8	8	8	64	4	256		4	8		64
MH025	0.25	0.25		0.12	0.12	0.5				0.5	0.5	0.13	32	0.5	256		0.5			4
MH026	0.25	0.25		0.12	0.12	0.5				1	1	0.25	16	2	256		2			8
MH027	0.25	0.25		0.12	0.12	0.5				0.5	0.5	0.12	16	1	256		1			4
MH028	0.25	0.25		0.12	0.12	0.5				0.5	0.5	0.25	16	0.5	256		0.5			4
MH029	16	0.25		1	2	1	8					8	64	8	256	2	8		2	16
MH031	0.5	0.25		1	2	8				8	8	0.5	64	4	256		4			64
MH032	0.25	0.25		1	2	1				8	8	0.5	64	4	256		4			16
MH033	1	0.25		0.12	0.12	1				1	1	1	16	1	256		1			8
MH034	0.25	0.25		0.12	0.12	1	1					0.12	32	0.5	256	1	0.5		1	4
MH035	16	0.25		1	2	2	8					8	16	8	256	16	8		16	16
MH038	0.25	0.25		0.12	0.12	0.5	1					0.25	16	0.5	256	1	0.5		1	4
MH039	0.25	0.25		0.12	0.12	0.5	1					0.12	32	0.5	256	1	0.5		1	4
MH043	0.25	0.25		0.12	0.12	0.5	1					0.25	32	1	256	1	1		1	8
MH044	0.25	0.25		0.12	0.12	0.5	1					0.12	16	1	256	1	1		1	8
MH046	0.25	0.25		0.12	0.12	0.5	1					0.25	16	1	256	1	1		1	4
MH047	0.25	0.25		0.12	0.12	0.5	1					0.5	16	0.5	256	1	0.5		1	2
MH048	0.25	0.25		0.12	0.12	1	1					0.12	32	1	256	1	1		1	4
MH049	0.25	0.25		0.12	0.12	0.5	1					0.12	32	1	256	1	1		1	8
MH050	0.5	0.25		0.12	0.12	0.5	1					0.25	16	1	256	1	1		1	2
MH054	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	32	16	16
MH055	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	32	16	16
MH056	16	0.25	16	1	2	8	8	16	32			8	64	8	256	16	8	16	16	16

C.2. continued

Isolate	AMP	CEFT	CLIND	DANO	ENRO	FLOR	GAMY	GENT	NEO	OTET	PEN	SPEC	SULFA	TET	TIA	TILD	TILM	TMS	TULA	TYL
MH057	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	8	16	16
MH058	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	32	16	16
MH059	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	32	16	16
MH060	2	0.25	8	1	2	8	8	16	32			4	64	8	256	16	8	16	16	16
MH062	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	16	16	16
MH063	16	0.5	16	1	2	4	8	16	32			8	64	8	256	16	8	8	16	16
MH064	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	32	16	16
MH065	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	8	16	16
MH066	16	0.5	16	1	2	8	8	16	32			8	64	8	256	16	8	8	16	16
MH067	16	0.25	8	0.12	0.12	0.5	1	2	8			8	16	8	256	1	8	16	1	2
MH069	16	0.25	8	0.12	0.12	1	1	2	8			8	64	8	256	1	8	16	1	4
MH070	16	0.25	8	0.12	0.12	0.5	1	2	16			8	64	8	256	1	8	16	1	4
MH071	8	0.25	8	0.12	0.12	0.5	1	2	16			8	64	8	256	1	8	16	1	2
MH072	0.25	0.25	8	0.12	0.12	0.5	1	1	4			0.25	16	0.5	256	1	0.5	8	1	8
MH074	16	0.25	16	1	1	2	8	2	32			8	16	8	256	16	8	16	16	16
MH075	0.25	0.25	8	0.12	0.12	0.5	2	2	8			0.12	32	1	256	1	1	16	1	4
MH076	0.25	0.25	16	1	2	8	8	16	32			0.5	64	8	256	16	8	8	16	16
MH077	0.25	0.25	16	0.12	0.12	0.5	1	2	8			0.25	32	0.5	256	1	0.5	32	1	8
MH078	0.25	0.25	8	0.12	0.12	1	4	2	4			0.12	32	0.5	256	2	0.5	32	2	16
MH079	0.25	0.25	8	0.12	0.12	0.5	1	2	8			0.12	32	0.5	256	1	0.5	8	1	4
MH081	0.25	0.25	8	0.12	0.12	0.5	1	2	8			0.5	32	1	256	1	1	16	1	4
MH082	0.25	0.25	8	0.12	0.12	0.5	1	2	8			0.5	32	8	256	1	8	8	1	4
MH083	0.25	0.25	8	0.12	0.12	0.5	1	2	4			0.12	32	0.5	256	1	0.5	16	1	8
MH084	16	8	16	0.12	0.12	8	8	2	4			8	32	8	256	16	8	32	16	16

C.2. continued

Isolate	AMP	CEFT	CLIND	DANO	ENRO	FLOR	GAMY	GENT	NEO	OTET	PEN	SPEC	SULFA	TET	TIA	TILD	TILM	TMS	TULA	TYL
MH085	0.25	0.25	8	0.12	0.12	0.5	1	2	4			0.12	32	0.5	256	1	0.5	16	1	4
MH086	0.25	0.25	8	0.12	0.12	1	1	2	8			0.12	32	0.5	256	1	0.5	8	1	4
MH087	0.25	0.25	8	0.12	0.12	0.5	1	1	4			0.12	32	0.5	256	1	0.5	16	1	8
MH088	0.25	0.25	16	0.5	0.5	1	1	2	8			0.5	32	1	256	1	1	16	1	8

C.3. Sequence information for Rule alignments. Length of region refers to length of filtered alignment (Reference) or rule match (Region).

		Length of region	% similarity of study isolates	% Isolates with matches
Danofloxacin				
*Rule 1 - Presence	Region 1	36-37	97-100%	97%
	Reference	37		
**Rule 2 - Presence	Region 1	60-61	98-100%	97%
	Reference	61		
Rule 3 - Absence	Region 1	81	100%	98%
	Reference	81		
Rule 4 - Presence	Region 1	51-61	83-100%	100%
	Reference	61		
Enrofloxacin				
*Rule 1 - Presence	Region 1	36-37	97-100%	97%
	Reference	37		
**Rule 2 - Presence	Region 1	60-61	98-100%	97%
	Reference	61		
Rule 3 - Presence	Region 1	39-40	97-100%	100%
	Region 2	38-48	79-100%	100%
	Region 3	31-33	93-100%	100%
	Region 4	30-31	96-100%	100%
	Region 5	53-54	98-100%	100%
	Region 6	105-107	98-100%	100%
	Region 7	76-82	92-100%	100%
	Ref region 1	40		
	Ref region 2	48		

C.3. continued

Enrofloxacin		Length of region	% similarity of study isolates	% Isolates with matches
Rule 3 - Presence	Ref region 3	33		
	Ref region 4	31		
	Ref region 5	54		
	Ref region 6	107		
	Ref region 7	82		
Florfenicol				
*Rule 1 - Presence	Region 1	36-37	97-100%	97%
	Reference	37		
Rule 2 - Presence	Region 1	48-49	97-100%	97%
	Region 2	53-54	98-100%	97%
	Region 3	61	100%	96%
	Ref region 1	49		
	Ref region 2	54		
	Ref region 3	61		
Tetracycline				
Rule 1 - Presence	Region 1	60-61	98-100%	91%
	Reference	61		
Rule 2 - Absence	Region 1	683-684	99-100%	100%
	Region 2	401	91%	100%
	Ref region 1	684		
	Ref region 2	436		

C.3. continued

Tildipirosin				
Rule 1 - Presence	Region 1	50	100%	93%
	Region 2	130	98-100%	93%
	Region 3	30-37	81-100%	93%
	Region 4	127-157	80-100%	93%
	Region 5	127	100%	93%
	Ref region 1	50		
Rule 1 - Presence	Ref region 2	130		
	Ref region 3	37		
	Ref region 4	157		
	Ref region 5	127		
Rule 2 - Presence	Region 1	60-61	98-100%	93%
	Reference	61		
Tilmicosin				
*** Rule 1 - Absence	Region 1			100%
	Region 2			100%
	Reference			
Rule 2 - Absence	Region 1	57-61	93-100%	100%
	Reference	61		
Tulathromycin				
*Rule 1 - Presence	Region 1	36-37	97-100%	97%
	Reference	37		
**** Rule 2 - Absence	Region 1			
	Reference			

C.3. continued

Tulathromycin				
Rule 3 - Presence	Region 1	52-53	98-100%	96%
	Reference	53		

C.4. Annotations for Model rulesets. Functional annotation of region refers to BLAST annotation of filtered alignment (BLAST reference) or RAST annotation of isolate match (Region)

		Functional annotation of region	Interesting Surrounding genes	
Danofloxacin				
*Rule 1 - Presence	Region 1	between tetR and tetH/ANT(2``)-Ia	transposases, ANT genes, bla-OXA2, msreE and mphE, PFGI-1 cluster genes, metal-binding genes	often duplicated on a high number contig
	Reference	upstream of tetR	tetH, other ARGs	Mixture of *Pasteurellaceae*
**Rule 2 - Presence	Region 1	sequence before HigA/HigB (toxin-antitoxin genes)	HigA/HigB, permease, lipase, cell wall, ribosomal methyltransferase	Mostly found at beginning or end of large contigs (>100kb)
	Reference	upstream of transcriptional regulator	IS481 transposase, HipA	MajorityMH
Rule 3 - Absence	Region 1	sequence before transposase	transposase, integrase, plasmid stabilization protein, mobile element protein, antirestriction protein, phage proteins, sig70, LPS proteins	most had >10 copies of the sequence, associated with various CDS but always IS481-like transposase
	Reference	upstream of IS481 transposase	adhesin, transcriptional regulator, DNA bidning protein	Majority MH
Rule 4 - Presence	Region 1	DNA topoisomerase section	sialic acid sugar isomerase, Pst permease genes, phosphate regulon, PTS, lysine synth gene aspartokinase	Some genomes with SNP (a to g) in middle, one genome with several mismatches
	Reference	DNA topoisomerase section	pstSABC, Pho genes, PTS system genes	Majority MH
Enrofloxacin				
*Rule 1 - Presence	Region 1	between tetR and tetH/ANT(2``)-Ia	transposases, ANT genes, bla-OXA2, msreE and mphE, PFGI-1 cluster genes	often duplicated on a high number contig
	Reference	upstream of tetR	tetH, other ARGs	Mixture of *Pasteurellaceae*
**Rule 2 - Presence	Region 1	sequence before HigA/HigB (toxin-antitoxin genes)	HigA/HigB, permease, lipase, cell wall, ribosomal methyltransferase	Mostly found at beginning or end of large contigs (>100kb)
	Reference	upstream of transcriptional regulator	IS481 transposase, HipA	Majority MH

C.4. continued

		Functional annotation of region	Interesting Surrounding genes	
Enrofloxacin				
Rule 3 - Presence	Region 1	section of SurE	Each region within same area	Most had SNP (g to a) beginning third of match
	Region 2	Different sections of tRNA pseudouridine synthase TruD		
	Region 3			
	Region 4	between TruD and triose-phosphate isomerase		
	Region 5	beginning of triosephosphate isomerase		Most had SNP (c to t) middle of the match
	Region 6	end section of PTS mannose transporter subunit IID		Most had 2 SNP (a to g, c to t) first 1/3, last 1/3 of match
	Region 7	middle of PTS mannose transporter subunit IID		Most had 6 SNP across match
	Ref region 1	section of SurE	Each region within same area	Majority MH
	Ref region 2	Different sections of tRNA pseudouridine synthase TruD		
	Ref region 3			
	Ref region 4	between TruD and triose-phosphate isomerase		
	Ref region 5	section of triose-phosphate isomerase		
	Ref region 6	different sections of PTS mannose transporter subunit IID		
	Ref region 7			
Florfenicol				
*Rule 1 - Presence	Region 1	between tetR and tetH/ANT(2``)-Ia	transposases, ANT genes, bla-OXA2, msreE and mphE, PFGI-1 cluster genes	often duplicated on a high number contig
	Reference	upstream of tetR	tetH, other ARGs	Mixture of *Pasteurellaceae*

C.4 continued

		Functional annotation of region	Interesting Surrounding genes	
Florfenicol				
Rule 2 - Presence	Region 1	between tetR and tetH/ANT(2``)-Ia	transposases, ANT genes, bla-OXA2, msreE and mphE, PFGI-1 cluster genes	Had to use NCBI nucleotide sequence as no alignment could be made; several copies across genome; overlap between this and Rule1
	Region 2	between tetR and tetH/ANT(2``)-Ia		often duplicated with a gap mid-sequence
	Region 3	section of tetH		
	Ref region 1	upstream of tetR	other ARGs	Mixture of *Pasteurellaceae*
	Ref region 2	upstream of tetR	tetH, other ARGs	
	Ref region 3	section of tetH	tetR, other ARGs	
Tetracycline				
Rule 1 - Presence	Region 1	between hyp protein and exodeoxyribonuclease I or another hyp protein	several hyp proteins, Na/Ala symport gene, TriC cycle genes or dna repair, phage genes, and iron transport	duplicated
	Reference	hyp protein	RNase P RNA component class A, hyp proteins	Majority MH
Rule 2 - Absence	Region 1	ISSod 13 transposase	23S methyltransferase, FtsH, pyruvate dehydrog complex, various permeases	Had to use NCBI nucleotide sequence as no alignment could be made; several copies across
	Region 2	IS1595 trasnposase	23S methyltransferase, FtsH, pyruvate dehydrog complex, various permeases	Had to use NCBI nucleotide sequence as no alignment could be made; several copies across
	Ref region 1	many copies of region associated with IS481-transposase	adhesins, porins, various rRNA and hyp proteins	Majority MH
	Ref region 2	many copies of region associated with IS1595-transposase	adhesins, porins, various rRNA and hyp proteins	Majority MH

C.4 continued

		Functional annotation of region	Interesting Surrounding genes	
Tildipirosin				
Rule 1 - Presence	Region 1	iron transporter permease protein	First four regions within same contig containing iron transport genes, ribosomal proteins, OsmB, RTX toxin genes, macrophage infectivity potentiatior, PFGI-1 cluster	some had partial matches to regions with no tetR but tetH or ANT(2``)-Ia
	Region 2	upstream and beginning of tetR		
	Region 3	between tetR and tetH/ANT(2``)-Ia		mobile element proteins were often partial genes;
	Region 4	ISL3-family transposase or generic "mobile element protein"		
Rule 1 - Presence	Region 5	beginning or end of contigs	xylulose/xylose utilization, anaerobic carbon utilization genes, lysine biosynth gene or transposase and hyp proteins	Multiple copies, one typically matched contig of previous regions
	Ref region 1	upstream of iron ABC transporter permease	hyp prot	Mixture of *Pasteurellaceae*
	Ref region 2	upstream of tetR	other ARGs	
	Ref region 3	between tetR and tetH		
	Ref region 4	downstream of tetH and overlapping ISL3 transposase	tetH	
	Ref region 5	between hyp protein and ssDNA binding protein		
Rule 2 - Presence	Region 1		hyp protein	Only one CDS on contigs
	Reference	between hyp protein and DDE		Majority MH
Tilmicosin				
Rule 1 - Absence	Region 1	parA,parB, replicase, tRNA-Leu	plasmid replication and partitioning genes	Overlapping sections on same contig
	Region 2			
	Reference	tRNA-Leu, parA, replicase, parB	various regions spanning genes	Majority MH

C.4 continued

		Functional annotation of region	Interesting Surrounding genes	
Tilmicosin				
Rule 2 - Absence	Region 1	phage tail tip host specificity J	other phage tail tip genes, phage tail genes	Often multiple, some as only CDS on contig. Sometimes contained a 3base gap and SNP (a to g)
	Reference	incomplete host specificity	phage genes	Majority MH
Tulathromycin				
*Rule 1 - Presence	Region 1	between tetR and tetH/ANT(2``)-Ia	transposases, ANT genes, bla-OXA2, msreE and mphE, PFGI-1 cluster genes	often duplicated on a high number contig
	Reference	upstream of tetR	tetH, other ARGs	Mixture of *Pasteurellaceae*
Rule 2 - Absence	Region 1	Need to find a way to use the model. It matches to many different regions across NCBI genomes		
	Reference	many copies of region associated with IS481-transposase	adhesins, porins, various rRNA and hyp proteins	Majority MH
		Functional annotation of region	Interesting Surrounding genes	
Tulathromycin				
Rule 3 - Presence	Region 1	upstream and beginning of Laccase	CopG, NrtY, DNA topoisomerase III	
	Reference	overlapping metal-binding and bilirubin oxidase		Mixture of *Pasteurellaceae*

Table C.5 Isolates with multiple errors. AB = antibiotic with error; MetaClass = AST interpretation used in the creation of the metadata file for the specific antibiotic; Train_Test = whether the error was in the training or the testing group during model creation; DANO = danofloxacin; ENRO = enrofloxacin; FLOR = florfenicol; TET = tetracycline; TILD = tildipirosin; TILM = tilmicosin; TULA = tulathromycin

Isolate	AB	MetaClass	Train_Test
MH012	FLOR	R	Train
	TET	R	Train
MH019	DANO	S	Train
	ENRO	S	Test
	FLOR	S	Test
	TET	R	Train
MH029	FLOR	S	Train
	TULA	S	Train
MH032	FLOR	S	Test
	TILM	S	Train
MH082	DANO	S	Test
	ENRO	S	Train
	FLOR	S	Train
	TET	R	Train
	TULA	S	Train
Isolate	AB	MetaClass	Train_Test
MH084	DANO	S	Test
	FLOR	R	Test
	TET	R	Train
	TILD	S	Train
MH088	DANO	R	Test
	ENRO	R	Train
MH2190	DANO	R	Train
	ENRO	R	Train
MH2597	DANO	S	Test
	ENRO	S	Test
	TILM	S	Train