DISSERTATION

EPIDEMIOLOGICAL INVESTIGATION OF ANTIMICROBIAL RESISTANCE IN BEEF PRODUCTION USING METAGENOMIC SEQUENCING

Submitted by

Enrique Doster

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2019

Doctoral Committee:

Advisor: Edward A. Hoover
Co-advisor: Paul S. Morley

Keith E. Belk
Zaid Abdo
Sheryl P. Gow

ABSTRACT

EPIDEMIOLOGICAL INVESTIGATION OF ANTIMICROBIAL RESISTANCE IN BEEF PRODUCTION USING METAGENOMIC SEQUENCING

Globally, the emergence of antimicrobial resistance (AMR) resulting in treatment failure is recognized as a growing public health threat. Antimicrobial use practices used in beef production are thought to be a direct driver of increasing antimicrobial resistance in pathogens and the environment, in part due to the higher volumes of antimicrobial drug necessary to treat cattle weighing 10 times more than an average person. This has led policy makers and public health organizations to promote "judicious use" or outright ban of antimicrobial drugs in livestock production. Use of antimicrobials is unavoidable for the treatment of disease and we must therefore learn how we can best adjust our AMD use to reduce selection of AMR pathogens. However, outside of important indicator organisms and pathogens, little is known about how different antimicrobial drug use practices affect communities of microorganisms, or microbiomes, and the AMR gene determinants, or resistome, shared between pathogen and non-pathogens alike. With advances in high-throughput sequencing (HTS), we can perform culture-independent studies and gain a better understanding of how antimicrobial drug use practices in livestock production affect AMR epidemiology.

This dissertation consists of five studies that employ HTS to characterize the microbiome and resistome of samples with differing AMD exposure along the beef production line. Projects begin with a look into the short-term effects on the microbiome and resistome of feedlot cattle following treatment with a macrolide drug, tulathromycin, in the manuscript "Investigating

Effects of Tulathromycin Metaphylaxis on the Fecal Resistome and Microbiome of Commercial Feedlot Cattle Early in the Feeding Period". Fecal samples collected in this project also were processed with aerobic culture, polymerase chain reaction (PCR), and lateral flow immunoassay for identification of *Salmonella enterica* and the comparison of these results are presented in "A Cautionary Report for Pathogen Identification Using Shotgun Metagenomics; a Comparison to Aerobic Culture and Polymerase Chain Reaction for *Salmonella enterica* Identification". Samples collected as part of a longitudinal study in feedlot cattle were analyzed to characterize the associations between AMD use and AMR in two bacterial species. These archived samples are leveraged for a broader understanding of AMR dynamics by adding a community-level perspective to results from aerobic culture. Results in individual cattle are presented in "Antimicrobial Drug Use in Beef Feedlots; Effects on the Microbiome and Resistome Dynamics in Individual Cattle" and results at the pen-level in "Metagenomic Investigation of the Effects of Antimicrobial Drug Use Practices on the Microbiome and Resistome of Beef Feedlot Cattle". Finally, in "Metagenomic Characterization of the Microbiome and Resistome in Retail Ground Beef" we examined the end of the beef production line by comparing the microbiome and resistome of retail ground beef products from either conventional production systems or those labeled as "raised with antibiotics" (RWA).

The five studies presented in this dissertation each contribute to the collective understanding of how AMD use in livestock production system can affect the ecology of AMR in microbial communities. These projects are useful first steps in learning to manage AMR in beef production systems; encompassing a targeted look at the use of one type of AMD, characterizing the resistome dynamics in individual cattle and pens over time in a feedlot, a comparison of the resistome in ground beef products, and many other aspects of AMR epidemiology. The final

study, describing limits to incorporating HTS for pathogen identification, serves as a cautionary reminder that with new technologies come new challenges and that research must keep pace.

ACKNOWLEDGMENTS

The projects in this dissertation were made possible thanks to funding support from the National Cattlemen's Beef Association; United States Department of Agriculture; U.S. National Institutes of Health; National Cattlemen's Beef Association; Mortimer- Beef Today-Elanco; Pattridge Family; Academy of Veterinary Consultants; Kingman Integrated Livestock Management scholarship, and the International Livestock Forum Fellowship.

Thanks to JBS, Feedlot Health Management Services and everyone that collaborated in the Microbial Ecology Group to help me complete the work in this dissertation; Maggie Weinroth, Steven Lakin, Kevin Thomas, Chris Dean, Ken Jones, Christina Boucher, Jaime Ruiz, Gina Georgnaras, Jennifer Parker, Cory Wolfe, Kathryn Crone, and Ben Ouyang; the many undergraduate students that kept our laboratory faithfully processing samples for me to analyze; Raf Neves, Lindsey Linke and Roberta Magnuson for teaching me the ropes of how to work with DNA extraction and helping me with the first project I ever worked on at Colorado State University; Joe Strecker and Tyler Eike for server support; Sheryl Gow, Zaid Abdo, Keith Belk, Edward Hoover, and Paul Morley for their dedication as my graduate committee.

Thanks to the Colorado State University DVM/PhD program for accepting me among their ranks and providing a community of supportive individuals that vastly influenced my PhD experience including Ed Hoover and Sue VandeWoude our fearless team managers, the team captain and morale booster Justin Lee, and all the special people in my cohort. I'd like to thank Noelle Noyes for mentoring me through difficult decisions and helping me find my mentor and friend, Paul Morley. Paul has patiently guided me throughout this whole PhD process and his unwavering support and confidence in my abilities was instrumental to my success as a graduate

student. Marilee Morley rapidly adopted me into the Morley family and shared their love to make me feel at home. My family is a constant source of motivation and never lost faith that I could achieve my monumental goal of being the first to pursue a doctorate degree. I am especially grateful for my mom's sacrifice to move to a new country and devote her life to provide better opportunities for her children. Without her, I would not understand the meaning of hard work and determination. Finally, to my amazing wife, Katie, who has supported me every step of the way and my dog, Moose, for giving me the perfect excuse to get outside and away from the computer to enjoy Colorado.

PREFACE

At the risk of sounding overly dramatic, I believe microbiology is currently in the early stages of a paradigm shift in the scale of the transition from the miasma theory to the germ theory in the 19th century. Instead of focusing on just individual bacterial species, the field of microbiology is rapidly expanding thanks to advances in high-throughput sequencing to consider multiple microorganisms at the same time (metagenomics). I remember my mentor, Paul Morley, first explaining to me why he was so excited about using metagenomics as a new tool to study the epidemiology of antimicrobial resistance (AMR) in livestock production. Then, he offered that I "ride the wave" with his lab and contribute to cutting edge research. Unbeknownst to him I was ready to sign-on before he shared his analogy, but this phrase would become a mantra of sorts to help me through the struggles of graduate school. Before getting to enjoy riding the wave, one must first pick the right surfboard and be ready to paddle until your arms feel like they might fall off. Before getting to enjoy contributing to cutting edge research, I had to paddle my arms off to learn the bioinformatic skills and knowledge necessary to discern which is the best tool for the job.

My doctoral dissertation aims to contribute to the understanding of how antimicrobial drug use in beef production affects the microbiome and the profile of antimicrobial resistance genes present (the "resistome"). Antimicrobial resistance is a pressing public health threat and I am fortunate to have worked with multidisciplinary team to evaluate several aspects important to beef production. The entirety of the results from my dissertation only push the needle forward a very small increment, but I feel privileged to have worked alongside amazing people to add

toward humanity's understanding of antimicrobial resistance. I will continue on this mission and undoubtedly will enjoy the company of the friendships I made along the way.

TABLE OF CONTENTS

CHAPTER 1: LITERATURE REVIEW: USES AND LIMITATIONS OF HIGH-THROUGHPUT SEQUENCING TO UNDERSTAND THE EPIDEMIOLOGY OF ANTIMICROBIAL RESISTANCE AND PATHOGENS IN BEEF PRODUCTION

GLOBAL BURDEN OF ANTIMICROBIAL RESISTANCE (AMR)

Concurrent with the discovery of penicillin, a medical marvel in 1940, Abraham and Chain (1) found that some bacteria produce an enzyme that can destroy the penicillin molecule, thereby making the organism resistant to treatment. Since then, an "arms race" has existed between innovative drug discovery for treatment of human and animal infections and the equally rapid phenomenon of antimicrobial resistance (AMR) (2–4). Now, in the 21st century, AMR in pathogens has emerged as one of the greatest public health challenges, threatening to push back medical progress by making previously treatable infections resistant to even the strongest antimicrobial drugs (AMDs) (5–7). A major cause for concern is that despite this decreasing effectiveness of currently employed AMDs due to resistance, there also is a relative decrease in the development of novel AMDs due to decreased investment, increased regulation, and high risk of failure (8). The CDC estimates that each year in the United States at least 2 million people are infected with drug resistant pathogens and 23,000 people die because of these resistant infections (5). Alarmingly, a report by the O'Neil Commission predicted that by 2050, 10 million deaths will be attributed to AMR infections (9). As a result, many countries and public health entities recognize AMR as a global public health threat and have created plans for combating AMR (9–11). Understanding drivers of AMR and improving systems for AMR surveillance are two important priorities shared across most plans and will be considered in the context of beef production in this review.

Use of AMDs is commonly accepted to be driving the rise of AMR in humans because the use of certain of these drugs or classes of drugs can also extend to livestock production, pets, crops, and aquaculture (12). While the magnitude of the threat from AMD use in livestock production is still poorly understood, beef production is of interest due to high AMD sale numbers and particular AMD use practices like the mass treatment of animals to prevent disease (13–16). In the US, the Federal Drug Administration (FDA) implemented the Veterinary Feed Directive (VFD) Regulation to promote the judicious use of AMDs and to restrict the use of AMDs for production purposes including growth promotion and feed efficiency (17). However, understanding how the epidemiology of AMR will shift in response to changes in AMD use practices is challenging and until recent years, studies relied on describing phenotypic patterns of resistance in certain "indicator" bacterial species. This work provides necessary insight into certain groups of bacteria, such as coliforms and obligate pathogens, but the results are complicated due to the sharing of AMR gene determinants between pathogens and non-pathogens, both of which could be reservoirs of AMR (18–21). We lack the ecological perspective of resistance and may be missing important dynamics between the microorganisms in a community (microbiome) and the AMR genes they carry (resistome).

Advancements in high-throughput sequencing now allow for non-targeted sequencing of a random portion of all the DNA present in a sample to either sequence across multiple genomes and study community-level structure (metagenomics) or capture extensive information about a single organism's genome, known as whole-genome sequencing (WGS). High-throughput sequencing allows an unprecedented look at the microbial environment and with its rapidly decreasing sequencing cost, has emerged as a powerful new tool for epidemiological studies of AMR. This chapter provides an update on the current state of AMR epidemiology in beef

production, describes relevant advances in high-throughput sequencing tools (metagenomics and WGS), and explores the uses for, and limitations of, incorporating these tools into future studies of AMR epidemiology.

EPIDEMIOLOGICAL STUDY OF AMR IN BEEF PRODUCTION

There is a dearth of knowledge regarding AMR dynamics in beef production. The extent to which different countries employ AMR surveillance varies greatly and much of the focus of surveillance is on AMR risk in humans, not animals. Notable examples of countries that have developed robust AMR surveillance programs in beef production include Canada (22), Denmark(23), and Sweden (24). Comparably, the United States is the largest beef producer globally and has limited AMR surveillance in beef production systems with patterns primarily for *Salmonella enterica* and nontype-specific *Escherichia coli*, as well as a growing program for *Enterococcus* spp isolates available through data from the United States Department of Agriculture's (USDA) National Animal Health Monitoring System (NAHMS) (25).

Conducting AMR surveillance and identifying drivers of AMR in beef production environments is challenging for a variety of reasons. Past research has primarily focused on just a few "indicator" bacterial species which overlooks the ecological perspective of AMR in the larger microbial ecology (25–28). Often, zoonotic enteric pathogens such as *Campylobacter* and *Salmonella* or indicator species like *E. coli* and *Enterococcus* spp are selected because of their importance in human disease, their ability to be isolated and cultured from healthy animals, and the availability of guidelines for the classification of resistance (29). Typically, culture is employed with selective media to isolate individual bacterial species found in samples from feedlot settings. Once isolated, bacteria are tested for the phenotypic expression of resistance to specific AMDs using either broth microdilution or Kirby-Bauer disk diffusion. In practice, these

tests work by exposing bacterial cultures to AMDs and establishing minimum inhibitory concentrations (MIC) with broth microdilution or establishing MIC "equivalent" values by converting the zone of inhibition measured by disk diffusion. Then, to designate whether an isolate can be classified as "resistant" or "sensitive" to each drug, MICs are compared to established breakpoints created by standardization bodies such as the Clinical Laboratory Standards Institute (CLSI) and the European Committee on Antimicrobial Susceptibility Testing (EUCAST)(30, 31). In cases where bacterial species do not have established breakpoint for certain AMDs, laboratories must set their own therapeutically-relevant thresholds.

These methods provide a useful overview of general trends in resistance for important bacterial species but are accompanied by complicating factors that limit our understanding of AMR ecology and therefore our ability to manage resistance. Genes known to confer AMR when expressed in pathogens can be shared between bacterial species through horizontal gene transfer among a wide range of microbial hosts to be maintained in the environment (6, 32, 33). Isolating and testing just a few "indicator" species for resistance to certain AMDs, therefore, cannot capture the complexity and full potential for resistance in an environment. Additionally, research shows that the selection of two different indicator species can provide distinct answers about the effect of AMD use on resistance (34) and can be biased by laboratory methods employed (35, 36). For example, polymerase chain reaction (PCR) can be used to test a sample for the presence of multiple AMR genes; however, the creation of targeted probes for specific nucleotide sequences inherently limits analyses to identifying known AMR determinants. Therefore, AMR surveillance programs employing PCR must continuously be updated to include probe designs for AMR genes important to public health but could miss the presence of novel resistance genes and underestimate the risk of AMR. The metagenomic approach, made possible by high-

throughput sequencing, can be used to complement current methods for AMR research and surveillance by providing a non-targeted approach that represents the total resistance potential in a sample across all organism's present.

HIGH-THROUGHPUT SEQUENCING TECHNOLOGIES

Starting with the push to sequence the human genome, high-throughput sequencing (HTS) technologies improved rapidly in sequencing quality and capacity while also decreasing in cost (37, 38). These advances have spurred creation of the "omics" field of study that consist of non-targeted sequencing of a particular type of molecule in a sample; DNA (genomics), mRNA (transcriptomics), ancillary DNA characteristics such as methylation and folding that affect transcription dynamics (epigenetics), amino acids and proteins (proteomics), and predicted metabolic genes (metabolomics). As comprehensive reviews pertaining to the various types of HTS tools are available elsewhere, the focus of this chapter will be on two widely used genomic methods, metagenomics and whole-genome sequencing (39–41).

Metagenomic sequencing (MGS) is a holistic approach to comprehensively sample the DNA extracted from a community of microbes. Unlike PCR, which requires targeted probes for the amplification of select nucleotide fragments, metagenomic sequencing studies are commonly conducted in the following three ways: shotgun sequencing (or whole metagenomics sequencing), marker-gene amplicon sequencing (e.g. 16S rRNA for bacteria, internal transcribed spacer (ITS) for fungi), and environmental clone libraries (functional metagenomics). Metagenomic sequencing has been shown to work on a growing list of sample matrices, including soil, feces, water, meat, blood, air and even samples collected from surfaces using swabs. On the other hand, WGS entails using non-targeted HTS typically on a sample consisting of a single isolate to obtain detailed information about sequence variation along the entire

genome of that organism. WGS allows an unprecedented view at the genetics of a single bacterial species and allows finer scale ecological modeling.

Advantageously, HTS sequencing results and the bioinformatic methods used to analyze them are publicly available through the National Center for Biotechnology Information (NCBI) genetic sequence database service and the software repository website GitHub. The quantity of stored sequencing data is rapidly growing (42, 43) and provides a wealth of knowledge that can be re-analyzed as computational resources and analytic methods improve.

USES OF HTS TOOLS FOR EPIDEMIOLOGICAL STUDIES

High-throughput sequencing has enabled the fields of metagenomics and WGS which are uniquely suited to revolutionize the study of AMR in beef production by improving AMR surveillance and the identification of factors driving AMR (44–46). As with all new tools, the potential uses and limitations are still being tested and evaluated, but recent reports presented below highlight exciting ways to use HTS for the study of AMR epidemiology.

Metagenomics

The use of culture-independent metagenomic sequencing removes the need for AMR research to be targeted to just a few species. Notably, past research supports the "great plate count anomaly" which suggests that only 1% of all bacteria on Earth can be isolated using culture methods (47). Estimates refined a decade later further suggest that out of roughly 61 distinct bacterial phyla in soil, 31 cannot be cultured (48). Whereas culture-dependent methods overlook the role played by these "unculturable" bacteria, metagenomic sequencing captures their resistance potential in the community. This approach opens up the possibility of identifying known AMR genes, metabolic genes, and virulence factors all in the same sample (49–51). Further HTS can be used to study multiple mechanisms of resistance. For example, certain

organisms can become resistant to AMDs and also the antimicrobial properties found in metals (52–54). This is thought to occur either through co-selection when both genes are present and selected for in the same organism, or through a single mechanism, such as multi-drug efflux pumps, which can confer resistance to both the metal and the AMD. HTS can be used in this context to study the co-selection of multiple AMR gene determinants with metal resistance providing previously unknown insights into resistance mechanisms.

Studies using metagenomic sequencing to investigate drivers of AMR in beef feedlot production are contributing to the knowledge about how AMR occurs at distinct points throughout the production line. Bovine respiratory disease (BRD) is an economically important disease in North American beef cattle whereby cattle deemed at high risk for BRD are often given parenteral antimicrobials metaphylactically. Metagenomic studies found that the nasopharyngeal microbiome and resistome are significantly affected by the administration of an antimicrobial (55). However, this antimicrobial administration does not seem to significantly alter the fecal microbiome and resistome composition (56). Testing the impact of in-feed AMD use on the resistome in the rumen of feedlot cattle did not show significant differences between treated and non-treated groups (57). Another study characterized changes throughout the beef production system and identified a decrease in resistome diversity from when cattle arrived to the feedlot, to when they were shipped for slaughter, to when they were processed at the abattoir (56). By the time the final product was ready for retail packaging, the resistome was at an undetectable level (58). When looking at retail beef products, results suggest that AMD use has a minor effect on the overall resistome when comparing between conventional production system products and those labeled “raised without antimicrobials” (52).

We speculate that using metagenomics to further explore the feedlot environment and characterize shifts occurring in cattle entering and exiting the feedlot could identify management practices positively associated with reducing AMR. It is possible, for example, that beef feedlot environments could have small-scale areas of increased resistance, similar to patterns observed in the intensive care unit environment within a hospital (60). Studies show that man-made or built environment can have a strong effect on the microbiome and resistome (61–63) and offers the possibility that built environments in beef feedlots could be altered to manage AMR dynamics.

Whole-genome sequencing (WGS)

The application of WGS is particularly well suited for AMR research in beef production systems and has been the most widely applied of the "omics" tools in food safety research (64). Because a lot of AMR surveillance programs were already relying on isolating bacterial species using culture, WGS can easily be implemented into the workflow at the end of the typical protocol to further characterize specific isolates. The subtyping of pathogenic isolates is specifically necessary for epidemiological studies and until recently subtyping was performed using pulsed-field gel electrophoresis (PFGE) and multiple-locus variable tandem number repeat analysis (MLVA) (65, 66). Mounting evidence suggests WGS can help with outbreak investigations by providing increased discrimination of isolates into epidemiologically relevant groups. Thus, as of 2017, PulseNet International, the network of collaborating laboratories dedicated to surveillance for food-borne diseases, will transition away from PFGE subtyping and replacing their workflow with WGS using whole genome multilocus sequence typing (wgMLST).

The following examples provide evidence of where WGS has been successfully utilized. One study of enterohemorrhagic *Escherichia coli* (EHEC) O26 species demonstrated the capability of WGS to improve the overall identification of O26 strains while simultaneously screening for AMR and virulence genes (67). Bioinformatic strategies to characterize bacterial species in this detailed manner is being applied to a multitude of livestock species to help address knowledge gaps in AMR ecology. This work includes the ability to identify novel AMR genes in phenotypically resistant isolates of *Clostridium difficile* (68) and the discovery of plasmids encoding multi-drug resistance genes that are widely distributed in *E. coli* isolates from different livestock species (69). In another example, a study used WGS to characterize nontyphoidal *Salmonella* across the beef production system in Mexico and their results identified clonal expansion of typhoidal toxin genes and VirB/D4 plasmids containing a virulence factor (70). These data also demonstrated that more than two decades before the isolates identified had been previously implicated in human clinical cases of Salmonellosis in Mexico. These findings highlight the importance of clonal expansion and persistence of Salmonella in the beef production environment. Use of WGS for AMR surveillance and outbreak investigation is well established and is increasingly being adopted worldwide (46, 71–75), but further exploration is warranted to determine further potential uses.

LIMITATIONS AND CHALLENGES

Metagenomic sequencing and WGS are promising tools for the study of AMR epidemiology, but as with all tools, their use must be cautiously evaluated. A typical workflow for both metagenomics and WGS would include the same basic steps; sample collection, sample processing, DNA extraction, library preparation, sequencing, and subsequent bioinformatic analysis. There are limitations and challenges at each step, starting with the lack of standardized

laboratory methods for sample collection and processing. For example, evidence shows that the external validity of an experiment could be compromised by differences in sample handling which could significantly influence the microbiome composition (76, 77). Similarly, the choice of DNA extraction kit provides different results from the same sample (78–80). Further, the choice of kit for library preparation and the choice of sequencing platform can each bias the results contributing to the necessity of standardized protocols (81, 82). In samples with high background host DNA such as carcass trimmings and meat rinsates, the development of laboratory methods to improve "on-target" sequencing of AMR genes is necessary to minimize sequencing costs and optimize sequencing depth. Currently, two methods have been proposed; depletion of vertebrate host DNA using a methyl-CpG binding domain (MBD) (83) and an alternate approach using biotinylated baits to target-enrich AMR genes prior to sequencing (84). An important limitation is that currently no standard exists for estimating absolute quantities of the genetic features identified in each sample, but a multitude of methods are being developed to address this (85–87).

Once sequencing is completed, the choice of bioinformatic methods selected to analyze HTS results is a critical component that is under debate for both metagenomic sequencing (88–90) and WGS (74, 91, 92). Identifying which genes are present in samples is not as straightforward as using the Burrows-Wheeler Aligner (BWA) software (93) to map the sample reads to a reference database. This is of vital importance to AMR studies because the function of some AMR genes cannot be inferred simply by their presence in a sample. Some genes, like gyrA, encode for an integral part of cellular biology for bacterial species but can confer resistance to AMDs through a point mutation in the gyrA gene. This point mutation prevents fluoroquinolone antimicrobials from binding to their target, alpha-subunit, thereby making the

organism resistant. The typical method of alignment, usually with the bwa software(93) and a reference database would miss this nuance and could classify wild-type gyrA genes in a sample incorrectly to overestimate the presence of fluoroquinolone resistance. The additional step of confirming the location of the resistance conferring point mutation relies on extensively curated annotations and several groups are attempting to address this and improve AMR classification (94, 95). Tied to that, the validity of the results garnered from HTS relies heavily on the integrity, diversity, and accuracy of the reference databases employed for classification (49, 96, 97).

Finally, performing bioinformatic analysis using command line coding can be quite challenging and may discourage scientists from incorporating HTS into their workflow. The Galaxy platform (98) was created to make bioinformatic analyses more accessible to other researchers and their graphical user interface (GUI) can be run directly from a web browser which avoids installing any bioinformatic software. However, running larger experiments on the Galaxy platform can be slow and little support is provided to help guide analytic decisions. With continuing development of user-friendly bioinformatic applications and increasing awareness of the expertise a bioinformatician can impart to ensure that sequencing effort is optimized, HTS tools can overcome major limitations and contribute valuable data to our understanding of AMR dynamics.

CONCLUSIONS

Antimicrobial resistance is considered the quintessential One Health issue that intersects the health of people, animals, and the environment (99). Microbes in the environment contained AMR before people started using AMDs (2, 4) and growing evidence points to the importance of the environment as a reservoir for resistance (100–102). Tools using HTS will undoubtedly play a role in future epidemiologic studies of AMR ecology. For AMR surveillance, both

metagenomics and WGS are already starting to be used to contribute a holistic community-level perspective and a detailed characterization of AMR gene sequence variance in specific isolates, respectively. This use will continue to increase as sequencing costs decrease and bioinformatic pipelines are streamlined. Outside of adopting the latest technologies to study the association between AMD use and AMR, it is necessary to be able to accurately estimate the selection pressure faced by a microbiome exposed to AMDs. Though the first step toward that goal is access to AMD use records, most AMD estimates rely on sales data that are likely leading to biased estimates (103) since sales data does not equate to what the animal was actually exposed to. Comparing AMD use across different species or collection methods is difficult therefore establishing a method to standardize these exposures is an ongoing target of research (26, 34, 34, 104). It is likely that AMD is only one driver of AMR and future studies must comprehensively evaluate the impact of environmental variables that are often ignored in microbiome and resistome experiments (e.g. temperature, humidity, pH, etc.) as well as management variables (e.g. stocking density, AMD use records, weight, source, etc.). Tackling the complex challenges posed by AMR cannot be addressed by a single scientific field, just as there is not a single tool that can reduce the risk of AMR in isolation. Instead, it will require collaboration from multi-disciplinary research teams to bring different tools and perspectives to the research table.

CHAPTER 2: INVESTIGATING EFFECTS OF TULATHROMYCIN METAPHYLAXIS ON THE FECAL RESISTOME AND MICROBIOME OF COMMERCIAL FEEDLOT CATTLE EARLY IN THE FEEDING PERIOD

SUMMARY

Background

Use of metaphylaxis treatment in beef feedlot production systems is thought to increase antimicrobial resistance.

Objectives

Characterize changes in the resistome and microbiome associated with metaphylactic treatment with tulathromycin in feedlot cattle early in the feeding period.

Animals

Two pens of cattle were used, with all cattle in one pen receiving metaphylaxis treatment (800mg subcutaneous tulathromycin) at arrival to the feedlot, and all cattle in the other pen remaining unexposed to parenteral antibiotics throughout the study period.

Methods

Shotgun sequencing was performed on isolated metagenomic DNA, and reads were aligned to a resistance and a taxonomic database to identify alignments to antimicrobial resistance (AMR) gene accessions and microbiome content.

Results

Overall, we identified AMR genes accessions encompassing 9 classes of AMR drugs and encoding 24 unique AMR mechanisms. Statistical analysis was used to identify differences in the resistome and microbiome between the untreated and treated groups at arrival to the feedlot (Day

1) or at a second sampling time (Day 11), as well as over time. Based on composition and ordination analyses, the resistome and microbiome were not significantly different between the two groups at either time point. However, both the resistome and microbiome in both groups changed significantly between these two sampling dates.

Conclusions

These results indicate that the transition into the feedlot – and associated changes in diet, geography, conspecific exposure, and environment – may exert a greater influence over the fecal resistome and microbiome of feedlot cattle than common metaphylactic antimicrobial drug administration.

INTRODUCTION

One of the most critical periods in managing the health and wellbeing of beef cattle is when they are transitioned from less intensive production settings, such as pasture or backgrounding operations, to feedlots. During this transition, animals are exposed to varied stressors associated with handling, transport, processing, commingling, and a shift to a high-energy feedlot diet (105). In response to these stressors, animals may become more susceptible to infectious disease, such as those that cause bovine respiratory disease (BRD), the single largest cause of morbidity and mortality among feedlot cattle in the United States (106, 107). Because groups of cattle that are deemed to have a high risk for the development of BRD can already be incubating infections that can become life-threatening despite the absence of clinical signs, these cattle are sometimes administered antimicrobial drugs (AMDs) at the time they enter the feedlot, a practice that is known as metaphylaxis. Metaphylactic administration of an AMD to an entire group of cattle with a high risk of BRD can be highly efficacious in preventing life-threatening disease. Parenteral administration of therapeutic doses of tulathromycin have been shown to be

highly effective when used metaphylactically for preventing illness and death related to BRD (108–111). However, the administration of AMD's to groups of animals are subject to increasing concern and scrutiny because of the potential for public health impacts related to AMR in bacteria that may be transferred to consumers through the food chain or environmental routes.

In the most recent national survey data available, 45.3% of feedlots reporting metaphylaxis use of tulathromycin to prevent BRD when cattle arrived at the feedlot (112). Tulathromycin use has been demonstrated to be highly effective in reducing BRD morbidity in feedlot cattle with only minor adverse side-effects (15, 108). Cardiotoxicity has been reported, with the use of Tulathromycin, in several species such as mice and rabbits(113, 114). Tulathromycin is a macrolide, a class of antimicrobials considered critically important for human medicine (12). Despite increasing scrutiny of antimicrobial use practices in livestock production, little work has been performed to study the effect of tulathromycin metaphylaxis on antimicrobial resistance (AMR) in cattle. Past research has frequently focused on phenotypic resistance to a limited number of AMDs in one, or at most a few, bacterial species using traditional culture methods(115, 116). However, the response to antimicrobial use varies among bacteria and because resistance genes can be transmitted amongst a wide variety of bacteria; results found in one bacterial species cannot be extrapolated to the community level (117, 118). Such AMD exposures have the potential to affect the entire gut ecology, and as such, a broader perspective is needed in investigating potential effects of metaphylactic AMD use on microbial communities.

High-throughput sequencing (HTS) techniques now enable a culture-independent metagenomic approach that can be used to study the resistome and microbiome, allowing access to the complete repertoire of resistance genes and bacteria within a given sample. Therefore, this

study was conducted to investigate the impact of metaphylactic AMD use of tulathromycin on the fecal resistome and microbiome of commercial feedlot cattle in the early feeding period using shotgun metagenomics.

MATERIALS AND METHODS

Overview of study design and population

Two groups of cattle were identified for enrollment in the study before their arrival at a commercial cattle feedlot in Texas. Cattle were purchased from a single backgrounding facility and were delivered in two groups of 193 and 186 steers (300-400 kg body weight/animal) each. Each group was housed in separate pens after arrival (Day 1), and one group was randomly selected to be given parenteral tulathromycin metaphylactically while the other group served as an unmedicated control. All cattle in the group that had an AMD administered received a subcutaneous injection of 800 mg tulathromycin (Draxxin®; Zoetis, Florham, NJ) while cattle in the other group did not. This single tulathromycin exposure was expected to result in therapeutic tissue concentrations in the lung for up to 14 days (119), and this drug has a withdrawal period of 18 days in the U.S. with regard to slaughter for human consumption of tissues (120). Essentially all of this drug is eliminated unmetabolized from the body via biliary excretion and subsequent fecal elimination. With the exception of the administration of tulathromycin to the one group, both groups of cattle underwent identical arrival processing, including administration of vaccines for clostridial and respiratory diseases, avermectin anthelmintic, and application of growth-promoting hormone implants (Table 2.1). After initial processing and placement into pens, cattle were fed the same corn-based diet throughout 11-day study period which contained tylosin (also a macrolide class of antimicrobial) to prevent liver abscesses at an FDA approved target intake of 90 mg per head per day and ionophore feed additives (monensin) conforming to nutritional

recommendations of the National Research Council (121). Cattle were provided *ad libitum* access to water and their health and welfare were monitored daily by trained feedlot personnel under the supervision of consulting veterinarians.

Fecal samples were collected from cattle per rectum at arrival (Day 1), and 11 days later (Day 11). After transport to the laboratory, fecal samples were processed to isolate total metagenomic DNA, upon which shotgun metagenomic sequencing was performed. During the 11-day study period, no cattle were identified as being ill, and therefore none received additional therapeutic AMD treatments.

Sample collection

Using individually packaged sterile gloves a total of 379 fecal samples (≥ 25 g/sample) were obtained per rectum from each steer at arrival processing, prior to tulathromycin injection of all cattle in one group. Each fecal sample was placed into a sterile Whirl-Pak bag (Nasco). Fecal samples were then placed into coolers with ice packs and transported to the laboratory within 8 hours of sample collection for storage at -80°C.

As part of another study evaluating methods for *Salmonella enterica* identification, all samples were processed prior to freezer storage with aerobic culture and lateral-flow immunoassay strips. Three cattle were identified as culture-positive for S. enterica; these 3 animals were from the AMD administered group. Not all animals that were sampled on Day1 were re-sampled. Animals that were sampled at the second sampling time, Day 11, included the 3 animals that were positive for S. enterica and 31 randomly selected animals. These 34 cattle (17 per group) were again palpated per rectum with sterile gloves to collect feces. Four animals had minimal feces in the rectum at this time (2 per group); therefore, fecal samples were

collected from 30 cattle (15 per group) and transported on ice to the laboratory for frozen storage. Only calves that had fecal samples at both time points were selected for further genomic investigation and processed for shotgun metagenomic sequencing. A total of 60 fecal samples collected at the two time points (Day 1 and Day 11) were analyzed.

DNA extraction

The 60 fecal samples were thawed at room temperature and total DNA was isolated. To remove excess plant debris and decrease inhibitors in fecal DNA samples, 10 grams from each sample were mixed with 30 mL of buffered peptone water (BPW), vigorously shaken, and allowed to sediment for 10 minutes. Supernatant was transferred to sterile 50 ml conical tubes and centrifuged at 4,300 ×g for 10 min at 4°C. Resulting pellets were rinsed with 5 mL of molecular-grade 1X phosphate buffered saline (PBS) and centrifuged again (4,300 ×g, for 10 min, at 4°C). After removal of supernatant, total DNA was isolated from the pellet using the PowerMax Soil DNA Isolation Kit (MO BIO Laboratories) following the manufacturer's protocol. DNA concentration and quality were evaluated using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, Inc.). Using NanoDrop results, samples with 260nm:280nm ratios >1.3 and DNA concentrations >20 ng/µl were sent for sequencing; samples that did not meet the concentration threshold were concentrated by ethanol precipitation before sequencing.

DNA library preparation and sequencing

Purified DNA (100 µl aliquots) from all 60 samples were delivered to the Genomics and Microarray Core at University of Colorado Denver (Aurora, CO) for library preparation and sequencing. Genomic libraries were prepared using the TruSeq DNA Library Preparation Kit

(Illumina, Inc.) and next-generation sequencing was completed on the HiSeq 2000 (Illumina, Inc.) with 5 samples per lane, V4 chemistry, and paired-end reads of 125 bp in length.

Processing of metagenomic sequence data

De-multiplexed sequence reads were analyzed using the AmrPlusPlus bioinformatic pipeline (49). Starting with read trimming and quality filtering using Trimmomatic (122), AmrPlusPlus then identifies host DNA with alignment to the *Bos Taurus* genome (123) using the Burrows-Wheeler-Aligner (BWA) software (93) and removes those reads with SamTools (124) to create non-host reads for subsequent characterization of the resistome and microbiome.

Analysis of sequencing quality

The FastQC software (125) was used to assess sample read quality. Summary statistics regarding the number of raw, trimmed, and non-host reads for each sample were compared using generalized linear models with the "glm" function and the R platform (126) to assess systematic bias across the following sequencing metadata: sequencing run, batch, and lane. For study design metadata, primary comparisons of interest were between treated vs. untreated cattle, and between sampling time points (Day 1 versus Day 11). To test for potential DNA contamination, sample reads were aligned to the human genome using BWA and the number of successfully aligned reads in each sample were compared between groups using the "wilcox.test" function. Similarly, differences in sequencing results between sample groups were tested with the Wilcoxon signed-rank test when comparing paired values from the same animal (Day 1 to Day 11) and the Wilcoxon rank-sum test was employed when comparing animals at either time point.

Resistome: Identification of resistance genes in metagenomic sequence data

In order to identify reads matching to resistance genes in the 60 samples, reads were aligned with BWA to the database MEGARes (49), a non-redundant nucleotide database of publicly available AMR gene sequences. For descriptive and statistical analyses, only genes with >80% "gene fraction", defined as the percent of nucleotides in each AMR reference gene that aligned to at least one read, were considered to be present in a sample. All gene accessions in the MEGARes database have been classified into an acyclic taxonomic hierarchy (drug class, mechanism, and group).

Accessions in the MEGARes database that are known to cause resistance as a result of single nucleotide polymorphisms (SNPs) in genes otherwise not associated with resistance were evaluated by visualizing the BWA alignments with Integrative Genomics Viewer (127). Reads were confirmed to align to the resistant allele sequence with 100% peptide homology (to allow for silent nucleotide substitutions) across the middle 95% of the reference AMR gene. Genes identified in our samples and included in this post-processing verification step were: *parE*, *rpoB*, *phoP, phoQ*, *evgS, evgA*, *crp*, *evgA, envR*, *marA*, *cpxA, cpxR*, *ompF*, and *blaR*. Any alignments that failed this verification step were removed from downstream analyses, as those reads likely represented wild-type DNA sequences that do not confer resistance to antimicrobials. Additionally, critically important resistance determinants (when expressed in human disease-causing agents) were identified *a priori*: (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat, cfr). Alignments to these genes accessions were specifically searched for in all 60 samples.

Microbiome: Identification and classification of bacterial sequences

Kraken (version 1) (128) was used to assign taxonomic labels to quality trimmed, paired non-host reads. To employ NCBI's RefSeq "Release 86" from January 12, 2018 (96), we created a custom kraken database consisting of RefSeq bacterial and archaeal genomes classified as either "reference genome" or "representative genome" and all complete viral genomes in RefSeq. Based on the recommendation of kraken's developers, all low-complexity regions were masked using DUST (129). Additionally, plasmid sequences were extracted from the genomic files and assigned to the "unidentified plasmid" taxa number ID "45202" to increase the specificity of taxonomic read classification and account for the horizontal transfer of plasmids in microbial communities (see full script at https://github.com/colostatemeg/meglab-kraken-custom-db).

Statistical analysis

Statistical analyses of the resistome and microbiome were accomplished using R packages "metagenomeSeq" and "vegan" (130, 131). Sparsely represented resistome and microbiome features (genes and taxa, respectively), that were identified in fewer than 5% of samples, were removed from further analysis to reduce likelihood that these features would bias abundance comparisons (131). Two different methods were used to normalize resistome and microbiome feature counts. Resistome counts were normalized using an equation (18) that allows for AMR gene abundance to be expressed as "copy of AMR gene per copy of 16S-rRNA gene" by accounting for differences in sequence length of AMR genes and bacterial load in the samples. Alignment to the full Greengenes database (97) using BWA with default settings in a paired-end manner was employed to identify 16S sequences in all non-host reads. Subsequently,

the “AMR gene abundance” of each gene identified within a sample was calculated using the equation (18):

$$\text{AMR gene abundance}=\sum_{1}^{n} \frac{N_{\text{AMR-likesequence}} \times L_{\text{reads}}/L_{\text{AMRreferencesequence}}}{N_{\text{16Ssequence}} \times L_{\text{reads}}/L_{\text{16Ssequence}}}$$

with $N_{\text{AMR-like sequence}}$ as the number of alignments to one specific AMR gene sequence; L_{reads} as the sequence length of the Illumina reads (125 nt); $L_{\text{AMR reference sequence}}$ as the sequence length of the corresponding AMR gene sequence; $N_{\text{16S sequence}}$ as the number of alignments to 16S sequences; and $L_{\text{16S sequence}}$ as the average length of the 16S sequences in the Greengenes database (mean= 1401 nt).

Resistome data were analyzed at the class and mechanism levels to avoid biased diversity measures caused by differences in the scientific criteria used for identification and publication of new resistance genes for different drug classes at the “gene” level (132). Alternatively, numbers of reads that matched microbial taxa were normalized to account for sequencing depth using cumulative sum scaling (CSS) (131). Sparseness of count data called for using a default percentile of 0.5 for normalization based on published recommendations (131). Corresponding taxonomic lineage for each taxon in the microbiome was identified and alignments were summed to these 6 Linnaean taxonomic levels: phylum, class, order, family, genus, and species. In total, there were 6 count matrices for the microbiome, but to reduce the repetitive reporting of results at all levels and because results at lower taxonomic levels are not considered very reliable (133), statistical results for microbiome are presented at the phylum, class and order levels. In total, 8 unique normalized count matrices (i.e., 6 count matrices describing the microbiome and 2 count matrices characterizing the resistome) were analyzed and reported. Figures were created using the base plotting functions in R, the ggplot2 package, and the Tableau software (134).

Ordination generation and testing

Normalized count matrices were Hellinger-transformed (135) and used for ordination analysis with the metaMDS function from "vegan". The metaMDS function employs non-metric multidimensional scaling (NMDS) on Euclidian distances with random starts to discover a stable ordination solution for plotting on two dimensions. Significance of separation between sample groups was tested using analysis of similarities (ANOSIM) (136). To assess the degree of correlation between the resistome and microbiome, the "procrustes" function was used to superimpose metaMDS ordination graphs and minimize the sum of squared differences. In the same manner, the correlation between the untreated and treated group's microbiomes and resistomes were calculated at both Day 1 and Day 11. Then, the function "protest" was used to calculate a M^2 statistic for each procrustes result.

Richness and diversity comparisons

For all 8 count matrices, the richness (i.e., the total number of unique features in each sample) and Shannon's diversity (i.e., the number and proportion of unique features in each sample) were compared between sample groups using the "wilcox.test" function in R.

Analysis of log-fold change in abundance

In order to identify specific features in count matrices with significantly different numbers of alignments between sample groups, metagenomeSeq's "fitZig" function (131) was used to fit multivariate zero-inflated Gaussian mixture models for all 8 count matrices separately. To avoid spurious feature comparisons, only features present in abundances greater than the 15^{th} quantile in each count matrix were considered. Statistical models consisted of fixed effects for

sample group (e.g., treated vs. untreated, or Day 1 versus Day 11) and sequencing batch number. The option "useMixedModel" and "block" was employed to account for repeated measures on cattle. Pairwise comparisons of feature abundance between sample groups were evaluated using limma's "makeContrasts" function (137) on the multivariate model, using alpha=0.05 on adjusted *P*-values as the cut-off value for statistical significance. This function outputs an estimate of the $\log_2$-fold change in abundance between groups for each feature (i.e., class/mechanism/phylum/order/etc.) with an associated *P*-value adjusted for multiple comparisons using the Benjamini-Hochberg procedure (138).

Data Submission

Quality-trimmed sequencing reads for all 60 samples described in this project have been deposited to the NCBI collection of biological data (BioProject). Accession PRJNA309291 ID: 309291

RESULTS

Sequencing results

Across all 60 samples, shotgun metagenomic sequencing generated 5.89 billion reads (2.95 billion paired reads) with an average of 49.1 million paired-end reads per sample (range 13.49 – 80.36 M, Supplemental Table 2.1). The average Phred quality score of raw reads across all samples was 35.2 (range 34.54 – 35.82) and only 4.4% of all reads were removed due to low quality (minimum per sample = 2.48%, maximum = 8.21%). Of the remaining reads, 19.69% (55.44M reads) were identified as bovine DNA and removed from subsequent analysis; 3 samples contained nearly 37% bovine DNA (probably because the feces were relatively scant in the rectum of these cattle at the time of sampling) and the other 57 samples ranged from 19.69%

to 27.11%. Alignment of non-host reads to the human genome identified on average 991,958 reads per sample (range = 210,246 – 4,639,154) and suggested minimal sample contamination (2.6% of reads across all 60 samples). There was a small, statistically significant difference in Phred scores when comparing samples by time and treatment due to high quality reads in all 60 samples (mean = 35.23, range = 34.54 – 35.82). This difference was not considered to be biologically meaningful. Additionally, because no other metadata comparisons yielded statistically significant differences, our results suggested that there was no systematic bias in sequencing effort.

Resistome composition

4,054,637 reads aligned to 208 AMR gene accessions in the MEGARes reference database. Following confirmation of genes conferring resistance due to single nucleotide polymorphisms (SNPs) and removal of sparsely represented genes (i.e., those found in less than 3 samples), there were 134 unique gene accessions in the MEGARes database that were identified from 3,773,873 reads. In all, these represented resistance to 9 unique AMR drug classes via 24 mechanisms of resistance, though the clear majority of reads aligned to gene accessions that confer resistance to tetracycline and the macrolide-lincosamide-streptogramin (MLS) class of antibiotics (76% and 18% of aligned reads, respectively). More than 99% of reads that aligned to tetracycline resistance gene accessions are known to confer resistance through ribosomal protection proteins, and 77% of the reads that aligned to MLS resistance gene accessions are known to confer resistance through macrolide efflux pumps. The 7gene accessions associated with multi-drug resistance (e.g., non-specific multi-drug efflux pumps) and resistance to the following drug classes; phenicol, bacitracin, fluoroquinolones, cationic antimicrobial

peptides, aminoglycosides, and betalactams. This pattern of fecal resistome composition was observed in both study groups and was seen in samples collected at both Day 1 and Day 11 (Figure 2.1). Of the *a priori* identified critically important resistance determinants, we only identified one AMR gene accession, bla(CTX-M), in a single sample from the treated group on Day 11.

The overall resistome composition was similar between the treated and untreated groups at both Day 1 and Day 11 (Figure 2.2). Apart from alignments to tetracyclines and MLS gene accessions, less than 3% of the resistome was characterized by alignments to multi-drug, betalactam, and aminoglycoside resistance gene accessions, with alignments to remaining classes of drugs each accounting for less than 1% of all alignments. While we did identify a difference in AMR Shannon's diversity when comparing treated and untreated cattle at Day 1 ($P = 0.05$), there was no evidence of significant differences in the relative abundances of AMR classes or mechanisms. In contrast by Day 11, the untreated group had significantly different AMR richness at the mechanism level ($P = 0.02$) and contained significantly higher abundance for the AMR mechanism, Tetracycline inactivation enzymes, than the treated group ($P < 0.05$).

In contrast to the lack of difference between treated and untreated groups either time point, there was a dramatic change in the resistome of both groups overtime between Day 1 and Day 11, such that there appeared to be a convergence toward a "common" resistome between groups. The untreated group's resistome shifted significantly at the class (ANOSIM $R = 0.22$, $P = 0.002$) and mechanism levels (ANOSIM $R = 0.30$, $P = 0.001$), as did the resistome of the treated cattle (ANOSIM $R = 0.21$, $P = 0.001$ for AMR drug class and ANOSIM $R = 0.40$, $P=0.001$ for AMR mechanism)(Figure 2.2). In both study groups, total AMR abundance, defined as "copies of alignments to AMR gene accessions per copy of 16S-rRNA gene", increased over

time (Day 1 to Day 11) from 3.04 to 5.29 in the untreated group and from 3.71 to 5.56 in the treated group. Consequently, the relative abundance of alignments to the two most abundant AMR classes, tetracyclines and MLS, increased between Day 1 and Day 11 for both the treated and untreated groups ($P<0.05$). The untreated group's resistome increased in abundance in two additional AMR classes, aminoglycoside and betalactam resistance ($P < 0.05$) albeit without exposure to these drugs. Correspondingly, the untreated group's significant changes in abundance were all increases in relative abundance of alignments to 5 of 20 resistance mechanisms between Day 1 and Day 11 ($P < 0.05$). Alternatively, the treated group had 15 mechanisms with significant changes in abundance, but 10 of 15 mechanisms decreased in abundance over time (Figure 2.3). Three AMR mechanisms increased in relative abundance in both groups, including tetracycline resistance ribosomal protection proteins, macrolide resistance efflux pumps, and class A betalactamases. The other 2 AMR mechanisms that increased in abundance over time differed by treatment group; aminoglycoside O-phosphotransferases and aminoglycoside N-acetyltransferases in the treated group, compared to increases in alignments to tetracycline inactivation enzymes and chloramphenicol acetyltransferases in the untreated group. Shannon's diversity indices of the treated group at the mechanism level decreased significantly over time ($P=0.04$), whereas there were no significant changes in richness or Shannon's diversity in untreated group (Figure 2.4). During these shifts in the resistome over time, procrustes analysis suggests that class level AMR resistome composition of treated and untreated cattle became more similar as they were significantly correlated only at Day 11 ($M^2 = 0.71$, $P = 0.02$).

While major trends in the most abundant AMR features can be observed at the treatment group level, there was considerable variation in the presence of low abundance AMR mechanisms between animals (Supplemental Figure 2.1**).** Interestingly, the number of samples

with alignments to phenicol and glycopeptide AMR classes increased over time in both study groups, though differential abundance comparisons were not possible due to their low abundance and sparse representation across all 60 samples. There were no samples with phenicol resistance gene accessions at Day 1, but alignments were present in 8 of 15 cattle from each treatment group by Day 11. Similarly, no samples had alignments to glycopeptide resistance gene accessions at Day 1; however, by Day 11 glycopeptide class resistance genes were identified in 3 of 15 untreated animals.

Microbiome composition

On average, 96.14% of sample reads were not classified as bacteria, archaea, or viruses (range = 93.71 – 96.98%). Alignments to a total of 5,910 taxa were identified across the 60 samples. Sparsely represented taxa were removed prior to normalization such that a total of 5,383 unique taxa were included in subsequent analyses (comprising alignments attributed to 38 phyla, 74 classes, 170 orders, 384 families, 1,211 genera and 3,943 species). The majority of microbiome alignments were to bacteria; alignments to Firmicutes, Bacteroidetes, Proteobacteria and Actinobacteria were most common, accounting for 99.7% of the total normalized hit counts at the phylum level (37%, 24%, 18% and 15%, respectively). At the class level, Clostridia (29%), Bacteroidia (21%), Gammaproteobacteria (10%), and Coriobacteriia (7%) were the predominant classes to which alignments were classified, representing more than two thirds of normalized counts. Clostridiales (32%), Bacteroidales (21%), and Enterobacteriales (6%) were the most abundant taxa at the order level (Figure 2.5).

No significant differences in the overall microbiome were observed between treated and untreated groups at Day 1 (ANOSIM $P > 0.05$), and taxa were not differentially abundant at the

phyla, class, or order level after adjusting for multiple comparisons. Similarly, at Day 11, ordination comparisons showed no distinct separation of microbial communities between the treated and untreated groups (Figure 2.6), and relative abundance of microbiome features did not differ at the phyla, class, or order levels. Moreover, richness and Shannon's diversity did not differ significantly between groups at either Day 1 or Day 11 (Figure 2.7). Unlike the resistome, procrustes analysis did not identify significant correlations between the groups' microbiomes at either time point.

Despite evidence suggesting that both groups had similar fecal microbiomes at Day 1 and Day 11, the composition shifted significantly over time in the feedlot at all microbiome levels for both the untreated group (phylum level: ANOSIM R=0.51, P=0.001) and the treated group (phylum level: ANOSIM R=0.50, P=0.001). The major shift that occurred in the composition of both study groups' microbiomes between sampling dates was characterized by an increase in the proportion of Actinobacteria and Firmicutes, which together accounted for 58% of the untreated and 64% of the treated group's resistome at Day 11 compared to 51% and 45% at Day 1, respectively (Figure 2.5). In the treated group, 17 of 38 phyla show significant changes in abundance over time, although there were only shifts in 7 of 38 phyla in the untreated group. Both groups' microbiome significantly increased in relative abundance of Firmicutes and Actinobacteria phyla, combined with a decrease in relative abundance of Gemmatimonadetes, Euryarchaeota, Candidatus Saccharibacteria, and Candidatus Planctomycetes ($P < 0.05$). Of the remaining phyla with significant changes in the treated cattle, 4 of 10 taxa increased in relative abundance, while the other 6 phyla decreased in abundance ($P<0.05$). Notwithstanding the major changes in microbiome composition, neither richness nor Shannon's diversity measures changed over time in either group.

Relationships between the fecal resistome and microbiome

Procrustes analyses suggests no statistically significant correlations were present between the resistome and microbiome within treatment groups at either time point ($P>0.05$).

DISCUSSION

Results of this study suggest that parenteral metaphylactic treatment of cattle with tulathromycin had minimal, if any, detectable short-term impact on the fecal resistome and microbiome of commercially raised feedlot cattle when evaluated using shotgun metagenomic sequencing. This is important because of critical concerns about public health in relation to AMD use in food-producing animals and also because this is an important drug for treatment and control of life-threatening respiratory disease in feedlot cattle. This study was conducted in a commercial feedlot operation to improve the practical relevance of our findings, but this also introduces important limitations. USDA data suggests that over 70% of feedlot cattle in the U.S. receive low doses of tylosin, a macrolide drug, in-feed for prevention of liver abscesses (112). While tylosin exposure of all study cattle may have confounded our ability to independently investigate the effects of tulathromycin (a different macrolide drug), this study aims to characterize the effect of additional metaphylactic AMD use in the context of commercial feedlot cattle production systems. Likewise, other studies have described that parenteral treatment with a tetracycline drug (oxytetracycline) can cause discernible changes in AMR even when cattle are also exposed to another in-feed tetracycline AMD (chlortetracycline) (55, 139). Comparing fecal samples collected at Day 1 to those collected on Day 11 uncovered several notable changes in the resistome and microbiome, suggesting that the transition from backgrounding operations to concentrated feeding in a commercial feedlot is a critical time for influencing the microbial community of beef cattle. The ancient phenomena of AMR is not likely to be eliminated from

microbial communities in natural environments (2), so techniques used to manage food animal populations (e.g., AMD use, diet changes, prebiotics, probiotics) need be evaluated as a way to support animal health and productivity while reducing AMR prevalence and transmission (140, 141). This study provides an ecological perspective suggesting metaphylactic tulathromycin treatment may be employed without incurring drastic changes to the resistome and microbiome of feces from typical feedlot cattle.

Between treated and untreated groups, shifting abundance from Day 1 to Day 11 in resistome and microbiome features differed by treatment, but ultimately maintained a "common" composition and total AMR abundance comprised principally of relatively few, highly-abundant taxa. In particular, procrustes analysis for the correlation between the groups' resistome was only significant at Day 11. Further, the resistome and microbiome of treated and untreated groups were largely similar on Day 11, suggesting that other selective pressures besides tulathromycin metaphylaxis (e.g., common environmental exposures, exposure of all study cattle to in-feed tylosin) are potentially stronger influences on changes to the resistome and microbiome in cattle that have been newly introduced to the feedlot environment. Limited sample size of 15 animals per treatment group could limit statistical power to detect differences in the resistome and microbiome caused by tulathromycin exposure, but a search of the relevant literature indicated a lack of power calculation tools for shotgun metagenomic sequencing experiments. The difference in Shannon's diversity observed between treated and untreated cattle could have occurred because individual cattle randomization into the two pens was not logistically feasible due to constraints imposed by the feedlot operator. Specifically, to address logistical complexities in cattle production, the cattle in this study were shipped in two separate container trucks from the backgrounding facility, and these separate groups automatically became the treated and

untreated groups upon arrival in the feedlot, as they were housed in separate pens due to arrival processing considerations. Nevertheless, this study contributes an ecological perspective into the microbial communities of individual feedlot cattle and emphasizes the utility of studying the bacterial community in beef feedlot operations to better characterize AMR dynamics.

This study verifies past reports that tetracycline and MLS resistance is commonly identified in cattle environments (142–144). Consistent with our group's previous research, resistome composition was largely dominated by abundance of sequence alignments to two mechanisms of resistance, representing tetracycline (ribosomal protection proteins) and MLS (macrolide efflux pump) classes of resistance which accounted for >60% and >28% of resistance in the treated and untreated study groups (58, 145). It is notable that there were no other parenteral antimicrobial drug treatments because of illness in the study cattle prior to Day 11, including a lack of exposure to drugs commonly used to treat illness in feedlot cattle such as tetracyclines, betalactams, and fluoroquinolones. It is possible that this influenced the decrease in alignments to AMR gene accessions in samples from both groups that encode for resistance to drugs not used in the study, such as bacitracin and fluoroquinolone. Interestingly, glycopeptide drug use is prohibited in beef cattle in the U.S., and while resistance was not identified at Day 1, three animals in the untreated group contained alignments to glycopeptide gene accessions. While this study cannot confirm the function of these genes, these results suggest glycopeptide resistance may have been present in the feedlot environment and spread amongst feedlot cattle. Similarly, chloramphenicol resistance was not identified at Day 1 and despite study cattle not being exposed to chloramphenicol drugs, at Day 11 both groups of cattle had 8 of 15 animals with alignments to chloramphenicol resistance gene accessions.

For the microbiome, time in the feedlot from Day 1 to Day 11 was associated with significant shifts in the microbial population in both groups, though ultimately maintained similar composition that was dominated by Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria. Temporal changes in the microbiome of cattle acclimatizing to feedlot rearing have reported dramatic changes in the nasopharyngeal microbiota of beef cattle after arrival at a feedlot (146, 147). These shifts in the fecal microbiome might be expected given the changes cattle are experiencing after arrival to the feedlot. In the microbiome of both groups, for example, we detected an increase of typical carbohydrate-digesting bacteria such as Lactobacillales, along with an increase of organisms with diverse metabolic functions within the phyla Firmicutes from Day 1 to Day 11 (148–150). Notably, the exposure to tulathromycin might have caused the decrease in relative abundance to the Proteobacteria and Verrucomicrobia phyla in the treated group. Both phyla consist of gram negative bacteria not typically considered macrolide targets, but their decrease in relative abundance is associated with concurrent increases in Firmicutes as reported with exposure a different macrolide drug, azithromycin (151). This corroborates previously published data asserting that microbiome similarity between cattle is strongly driven by exposure to comparable management practices and/or the same geographic region (152).

Though we were not able to obtain information about the management of study cattle before arrival at the feedlot (i.e., source of cattle, diet, antimicrobial use, etc.), the lack of major differences in the resistome between groups at Day 1 might be attributed to rearing in the same backgrounding facility under similar husbandry practices immediately prior to being shipped to the feedlot. It is important to note that the lack of difference between treated and untreated groups either at Day 1 or Day 11 could also be explained by the high abundance of sequences

(>90% relative abundance) coding for resistance to tetracyclines and MLS making a “core” resistome which could potentially mask important differences in less abundant resistance genes (153). The pharmacokinetics of tulathromycin tissue concentration have been previously described (154), so the choice of 11 days between sampling points ensured that tulathromycin was still in therapeutic concentrations, but its influence on the fecal resistome and microbiome is undefined and future studies should consider time series sampling to capture temporo-dynamic changes in AMR ecology. Future research is needed to estimate the risk of different resistome compositions compared to our understanding from AMR patterns found in certain pathogens through traditional culture-based approaches. Additionally, while sequencing processes and bioinformatic analyses techniques continue to improve, we need broad collaboration from the scientific community to standardize AMR gene nomenclature and bioinformatic analysis so that results can be comparable across studies (132, 155).

Table 2.1. Products administered to study cattle at the time of arrival-processing (Day 1).

Product type	Commercial name	Manufacturer	Volume per animal	Additional information
Antimicrobial*	Draxxin	Zoetis	8 cc	Macrolide antimicrobial for treatment of cattle at high risk for bovine respiratory disease (BRD).
Anthelmintic	Noromectin	Norbrooks Labs	7 cc	Ivermectin parasiticide for the treatment and control of internal and external parasites of cattle.
Anthelmintic	Safeguard	Merck Animal Health	18 cc	For use in beef cattle for the removal and control of lung, stomach and intestine worms.
Vaccine	BoviAnthelmintic-Shield GOLD	Zoetis	2 cc	Protects cattle from infectious bovine rhinotracheitis (IBR) and bovine viral diarrhea (BVD).
Vaccine	Vision® 7	Merck Animal Health	2 cc	For use in healthy cattle as an aid in the preventing disease caused by Clostridium spp.
Steroid implant	Revalor-XS	Merck Animal Health	Implant	Trenbolone acetate and estradiol. It improves rate of gain and feed efficiency.

*Only the treated group received the antimicrobial treatment

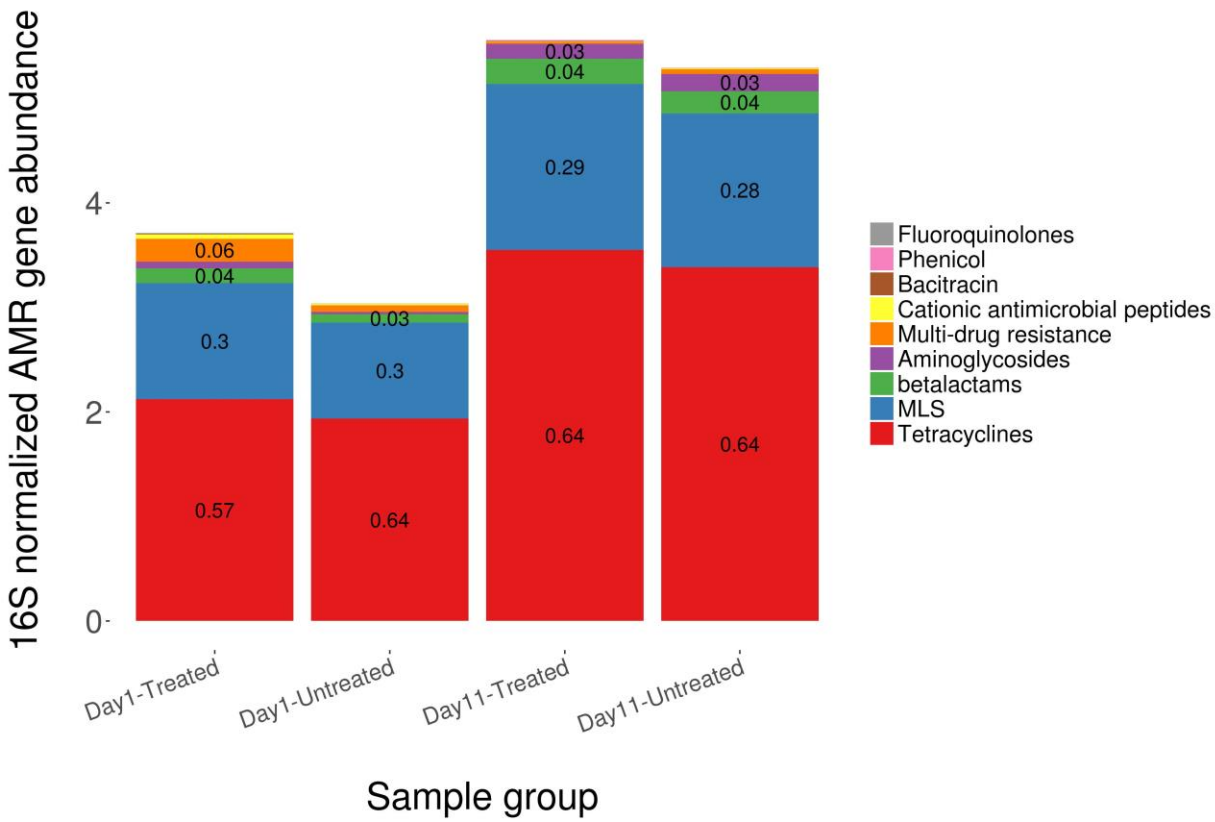


Figure 2.1. Total AMR gene abundance determined by shotgun metagenomic sequencing and normalized using 16S rRNA abundance, by drug class, among treated and untreated cattle in samples obtained at Day 1 and again at Day 11. Values are formulated from the number of reads that aligned to AMR genes and normalized to bacterial abundance characterized by alignments to 16S gene sequences from the Greengenes database.

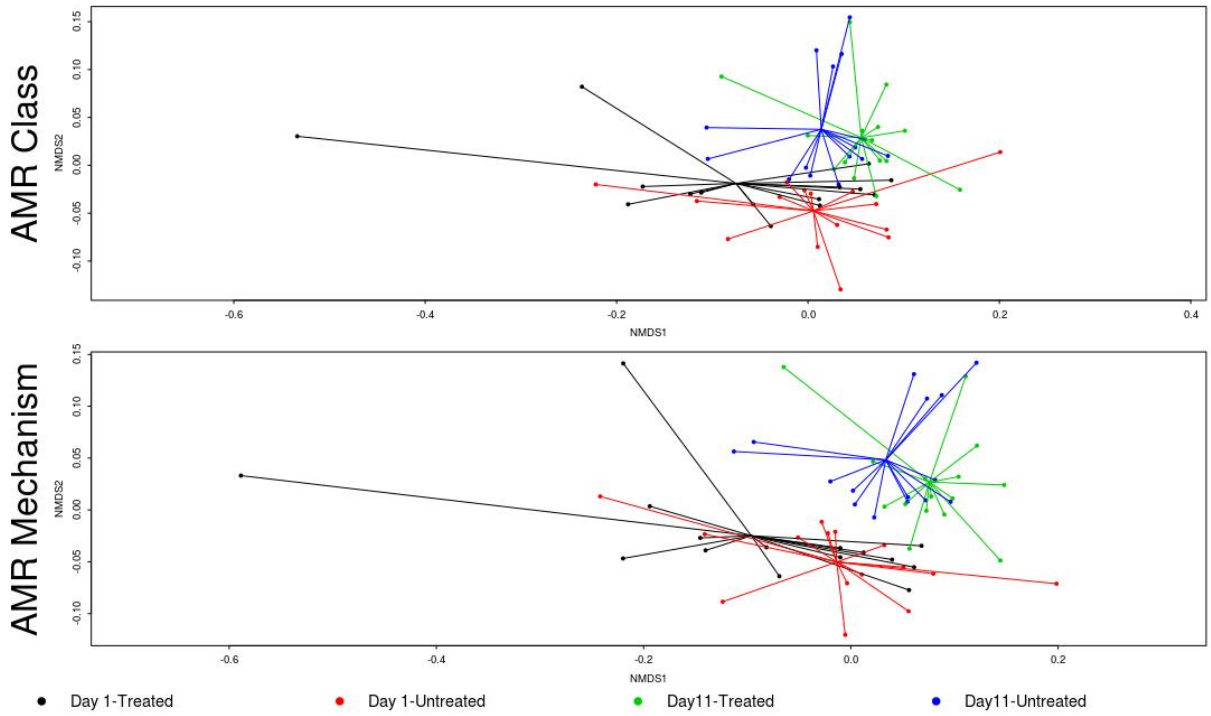


Figure 2.2. Ordination comparing resistome composition at the AMR drug class and resistance mechanism, using non-metric multidimensional scaling (NMDS), for the two study groups at Day 1 and Day 11. Separation of resistomes from treated and untreated cattle was not statistically significant at either Day 1 or Day 11 (Day 1 vs. Day 11; ANOSIM $P > 0.05$). However, the resistomes of the treated and untreated groups were significantly separated over time (Day 1 vs Day 11; ANOSIM $P < 0.05$).

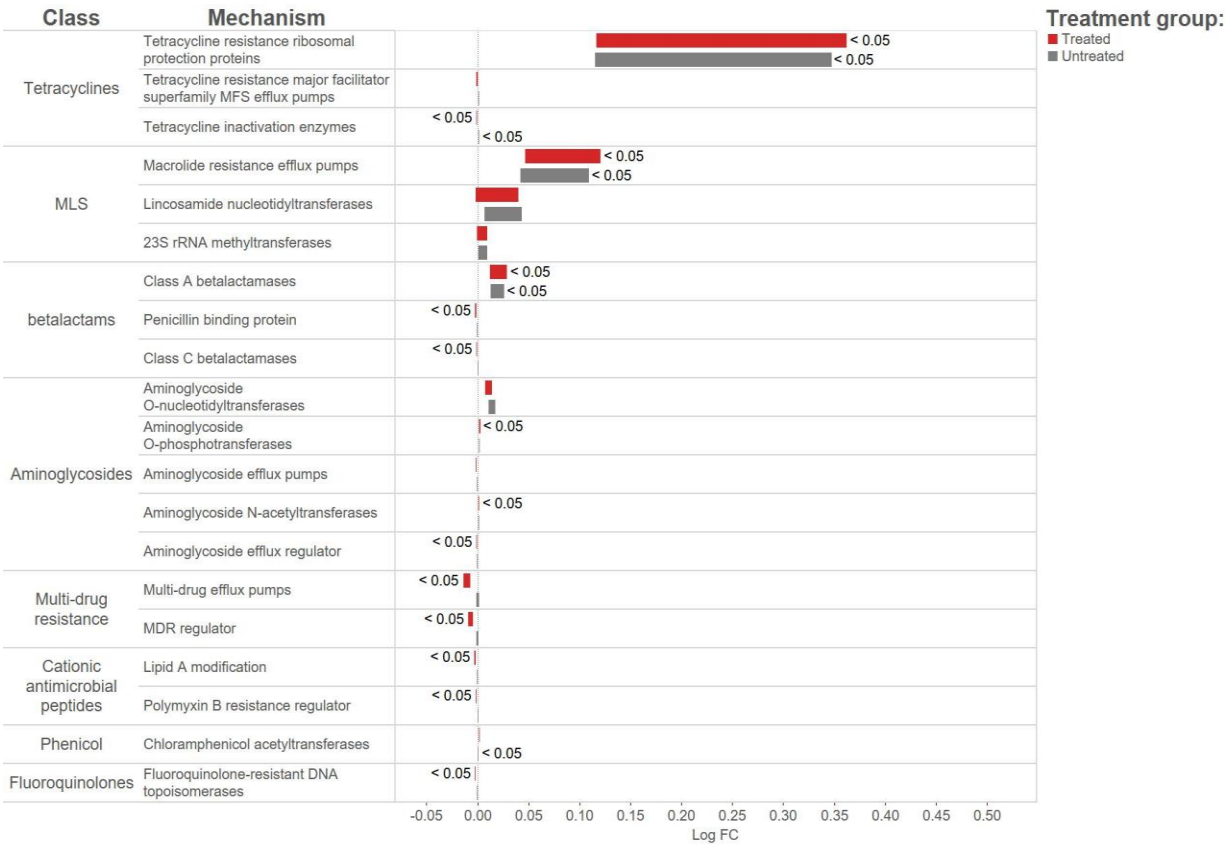


Figure 2.3. Log-fold change in abundance to AMR mechanisms for the treated (red bars) and untreated (grey bars) over time from Day 1 to day11. Bars to the right of the 0-line signify and increase in abundance, the size of the bars represent the average expression of the AMR mechanism and bars are labeled with adjusted p values < 0.05.

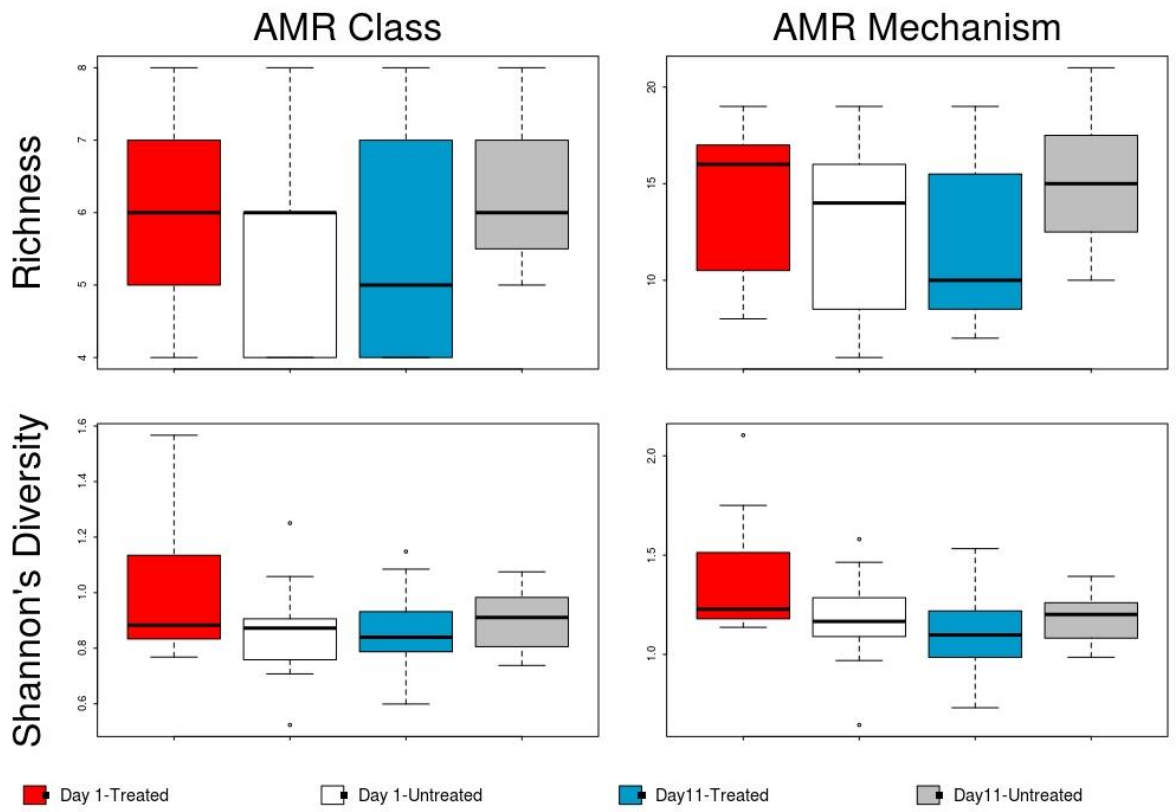


Figure 2.4. Boxplot of resistome richness and Shannon's diversity at the AMR class and mechanism levels of the two study groups at Day 1 and Day 11. The horizontal line is the median value, the middle box indicates the inter-quantile range, whiskers represent values within 1.5 IQR of the lower and upper quartiles, and individual points show outlier values.

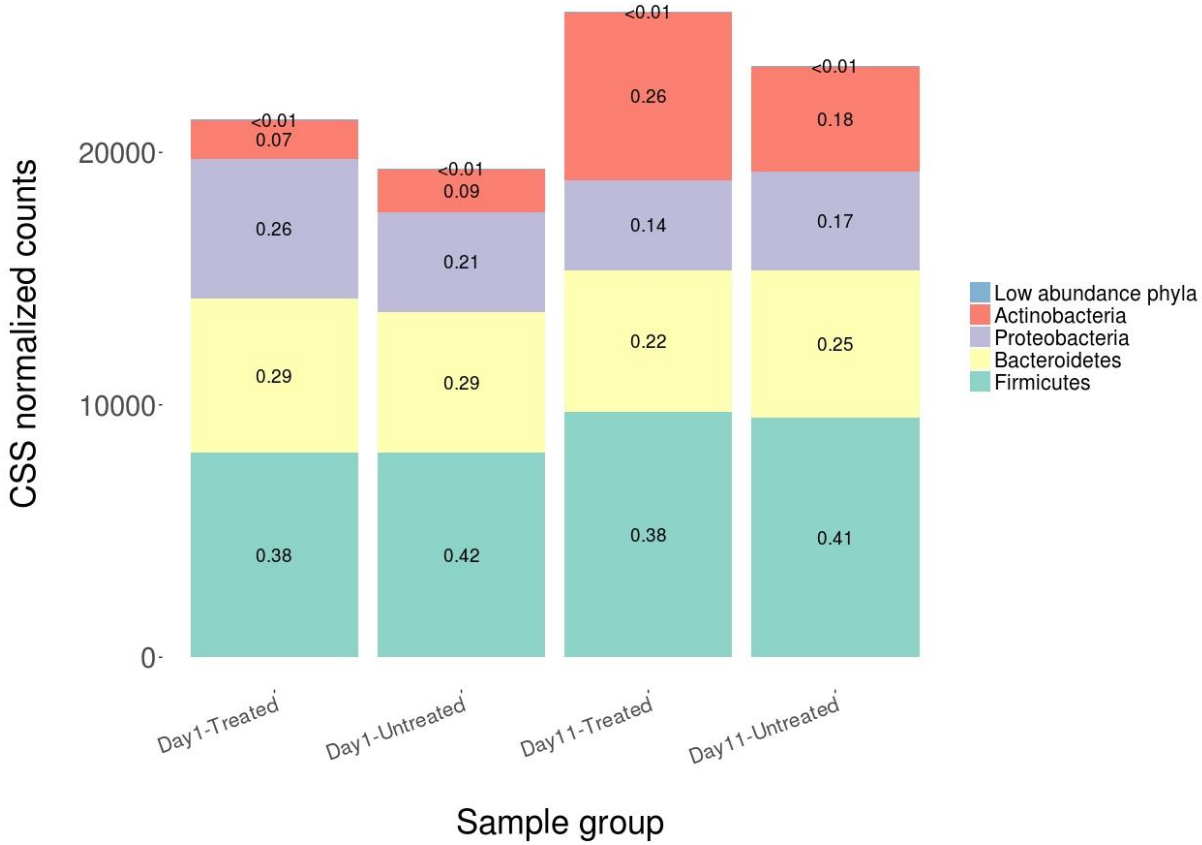


Figure 2.5. Average relative abundance of CSS normalized counts of shotgun metagenomic reads aligning to bacterial, archaeal and viral genomes at the phylum level for both study groups at Day 1 and Day 11. Phyla comprising less than 3% of each sample group were combined into the category “Low abundance phyla”.

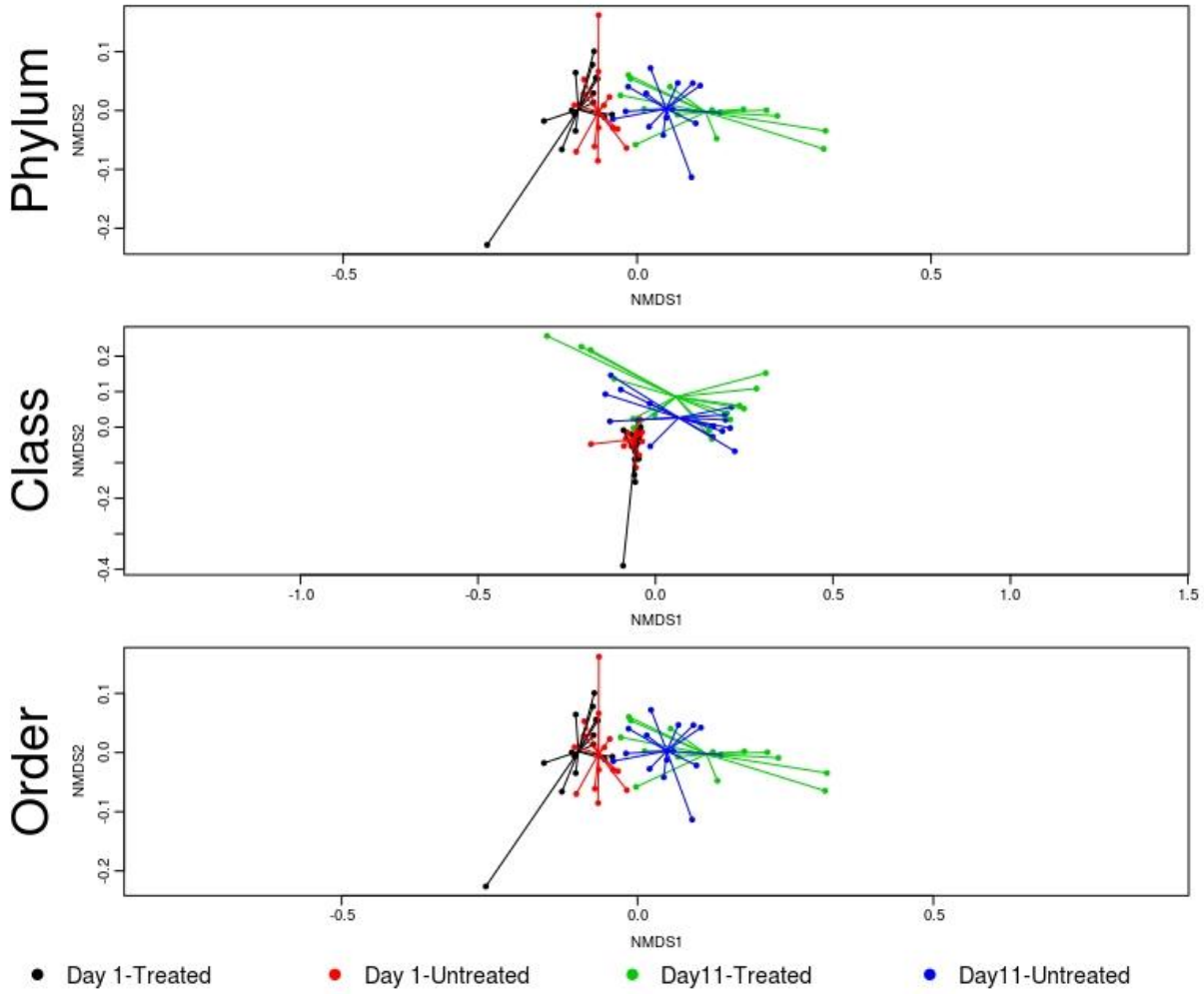


Figure 2.6. Ordination comparing microbiome composition at the phylum, class, and order levels, using non-metric multidimensional scaling (NMDS), for treated and untreated groups of cattle at Day 1 and Day 11. Separation of microbiomes from treated and untreated cattle was not statistically significant at either Day 1 or Day 11 (treated vs. untreated; ANOSIM $P > 0.05$). However, microbiomes for the study groups differed significantly over time (Day 1 vs Day 11; ANOSIM $P < 0.05$).

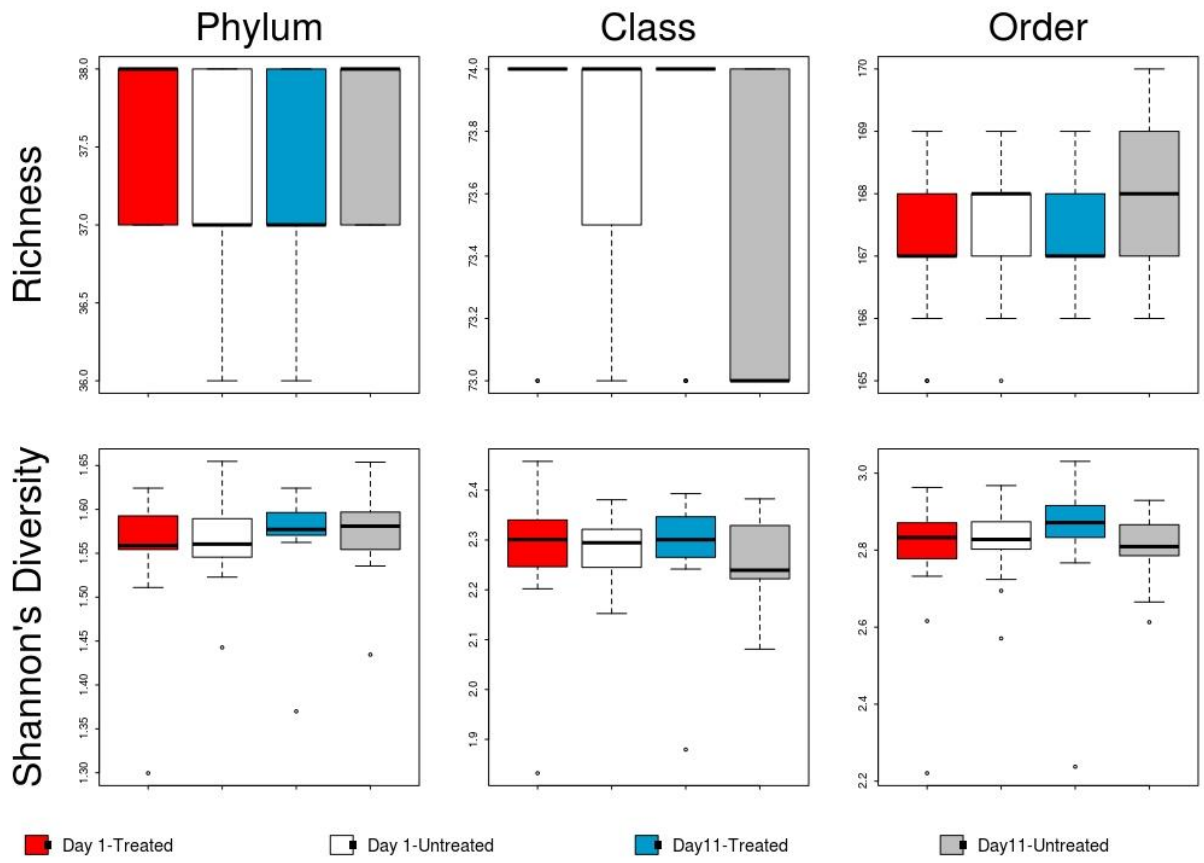


Figure 2.7. Boxplot of microbiome richness and Shannon's diversity at the phylum, class and order levels of the two study groups at Day 1 and Day 11. The horizontal line is the median value, the middle box indicates the inter-quantile range, whiskers represent values within 1.5 IQR of the lower and upper quartiles, and individual points show outlier values.

CHAPTER 3: A CAUTIONARY REPORT FOR PATHOGEN IDENTIFICATION USING SHOTGUN METAGENOMICS; A COMPARISON TO AEROBIC CULTURE AND POLYMERASE CHAIN REACTION FOR SALMONELLA ENTERICA IDENTIFICATION

SUMMARY

This study was conducted to compare aerobic culture, polymerase chain reaction (PCR), lateral flow immunoassay (LFI), and shotgun metagenomics for identification of *Salmonella enterica* in feedlot cattle feces. Fecal samples were collected from the rectum of 30 cattle upon arrival at a commercial feedlot and 11 days later. Samples were then analyzed in parallel using all 4 tests. Aerobic culture and PCR results had 100% agreement and indicated low abundance with 5.0% (3/60) of samples positive for *S. enterica*, which disagreed with results of LFI and metagenomic sequencing. Although low *S. enterica* prevalence restricted formal statistical comparisons, metagenomic analysis using k-mer alignment for classification using the RefSeq database identified that 18.3% (11/60) of samples contained reads matching to S. enterica. However further examination revealed that plasmid sequences are often included with bacterial genomes submitted to NCBI, which can lead to incorrect taxonomic classification. As such, we separated all plasmid sequences included with RefSeq genomes and reassigned them to a unique taxon representing plasmids. Reclassification of sequencing data using this revised database indicated that only 10% (6/60) of samples contained sequences specific for S. enterica, suggesting increased relative specificity. Reads identified using alignment were further evaluated using BLAST and NCBI's nr/nt database, which indicated that only 2/60 samples contained reads that exclusively aligned to *S. enterica* chromosomal genomes. This study re-affirms that the traditional techniques of aerobic culture and PCR provide similar results for *S. enterica*

identification in cattle feces. On the other hand, metagenomic results are largely influenced by the matching method and reference database employed meaning that computationally-based identification of bacterial species must be interpreted cautiously.

INTRODUCTION

The study and detection of microbial organisms has long been reliant on cultivation and characterization of certain species, but advancements in sequencing technologies have revealed an underlying microbial diversity largely ignored by culture-based techniques (156–158). High-throughput sequencing techniques now enable a culture-independent metagenomic approach that provides access to DNA from all bacteria (microbiome) within a given sample. This rapidly developing technology provides great potential for investigating the complexity of bacterial communities (19, 145, 159). However, there are limited numbers of investigations evaluating the relationship between metagenomic results and traditional diagnostic methods. Metagenomic approaches have been used to find novel pathogens when traditional methods were not fruitful (160, 161), but the increased sensitivity in metagenomic approaches raises important questions about their use in identification of foodborne pathogens in fecal samples. One such example is *Salmonella enterica*, an important zoonotic pathogen that causes over 93 million cases of gastroenteritis in humans globally every year (162) and has been implicated in outbreaks associated with beef products (163). Accurate identification and characterization of *S. enterica* is critical for improving food safety and preventing foodborne disease outbreaks. Therefore, utilizing samples obtained from another study of feedlot cattle (56), we compared a metagenomic approach for Salmonella enterica identification to the traditional techniques of aerobic culture, polymerase chain reaction (PCR), and lateral flow immunoassays (LFI).

METHODS

Sample collection

As previously described (56), 2 groups of cattle originating from a single facility were enrolled into the study before their arrival at a commercial cattle feedlot in Texas and individually sampled during arrival processing at the feedlot and 11 days later. Cattle were shipped, housed, and managed in two separate groups of 193 and 186 steers (300-400 kg body weight/animal). One group was selected to be administered a subcutaneous injection of 800 mg tulathromycin (Draxxin®; Zoetis, Florham, NJ), but otherwise the study cattle underwent identical management including administration of vaccines for clostridial and respiratory diseases, avermectin anthelmintic, and application of growth-promoting hormone implants (Table 3.1). Cattle were fed a corn-based diet with ionophore feed additives conforming to nutritional recommendations of the National Research Council (121), provided *ad libitum* access to water. Tylosin, a macrolide class of antimicrobial, was included in the feed to prevent liver abscesses at a target intake of 90 mg per head per day. Under the supervision of consulting veterinarians, trained feedlot personnel monitored cattle health and welfare daily.

On arrival processing, we attempted to sample 193 control and 186 treated cattle per rectum. Of these 379 cattle, 29 could not be sampled due to absence of feces in the rectum and another 4 samples had to be discarded for a total of 346 cattle sampled at arrival. In all, 29 cattle had less than 25 grams of feces for sample collection but were still included in the study. Three samples of those collected at arrival were culture-positive for *Salmonella enterica* (0.87% prevalence, 3/346); all three of these samples were obtained from cattle in the treated group. Fecal samples were stored in a sterile Whirl-Pak bag (Nasco), then placed into coolers with ice packs and transported to the laboratory. Within 8 hours of collection, samples were processed for culture and PCR and the remaining portion was placed in frozen storage (-80°C). Budgetary

limitations for shotgun metagenomic sequencing and PCR testing did not allow evaluation of all cattle sampled at arrival. Fecal samples collected on day 11 from 30 animals (15 treated and 15 untreated cattle) were selected, along with their arrival samples, for total DNA extraction and shotgun metagenomic sequencing. At day 11 sampling, we attempted to re-sample 17 cattle per rectum from each group. Of these 34 cattle, 4 could not be sampled due to absence of feces (including a steer that was culture-positive for *S. enterica* at arrival processing), and 3 samples had less than 25 grams of feces.

Salmonella enterica culture and PCR

As some cattle had scant feces present in their rectum at the time of sampling, samples of < 25 gm were processed in smaller quantities for aerobic culture. Upon delivery to the laboratory, all fecal samples were thoroughly mixed and aliquots of 4 gm of feces were removed from samples of >25 gm while 1 gm aliquots of feces were removed from smaller volume samples. The remainder of fecal samples were frozen at -80° C. Fecal samples undergoing aerobic culture were enriched for 18hrs at 43° C in tetrathionate broth (9:1 broth volume:fecal mass; Difco Laboratories, Sparks, MD). After initial enrichment, samples were mixed and 0.1 ml was passed into 10 ml of Rappaport-Vassiliadis R10 broth (Difco Laboratories) and incubated for 18 hrs at 43° C. After secondary enrichment, samples were thoroughly stirred and streaked for isolation on xylose-lysine-tergitol (XLT-4) agar plates (BD Diagnostic Systems, Sparks, MD). Bacterial growth was evaluated after 24 hours of incubation at 43° C. For samples that had colonies with characteristics indicative of *S. enterica* (smooth, round, black colonies), a single colony was arbitrarily chosen and streaked for isolation on tryptic soy agar with 5% sheep's blood (TSA) (BBL, Sparks, MD), and incubated for 24 hrs at 37C. Individual colonies were then

tested with polyvalent O antiserum (Difco Laboratories) for Salmonella serogroup confirmation. Antisera (Difco Laboratories) specific to 5 different common serogroups (B, K, D, C1, E) were used to further characterize the isolates. Additionally, for each sample during aerobic culture, lateral flow immunoassays strips (Reveal 2.0, NEOGEN Lansing, MI) were used to test the tetrathionate broth for *S. enterica* identification. These strips have been tested with horse fecal samples and show promising sensitivity and specificity for rapid identification of *S. enterica* (164, 165), but have not been tested extensively in cattle feces. Fecal samples were also tested by qPCR for *S. enterica* detection (Applied Biosystems, Foster City, CA) using a 2 ml aliquot of enriched culture media (TET). The target for this PCR is proprietary information, but this commercial PCR kit is routinely used at the Colorado State University (CSU) Diagnostic Medicine Center.

DNA extraction

Fecal samples selected for metagenomic sequencing were thawed and processed for DNA extraction. To remove excess plant debris and decrease inhibitors in fecal DNA samples, 10 gm from each sample were mixed with 30 mL of buffered peptone water (BPW), vigorously shaken, and allowed to sediment for 10 minutes. Supernatant was transferred to sterile 50 ml conical tubes and centrifuged at 4,300 ×g for 10 min at 4°C. Resulting pellets were rinsed with 5 mL of molecular-grade 1X phosphate buffered saline (PBS) and centrifuged again (4,300 ×g, for 10 min, at 4°C). After removal of supernatant, following the manufacturer's protocol total DNA was extracted from the pellet using the PowerMax Soil DNA Isolation Kit (MO BIO Laboratories). DNA concentration and quality were evaluated using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, Inc.). Samples with 260nm:280nm ratios >1.3 and DNA concentrations >20

ng/µl were sent for sequencing; samples that did not meet these thresholds were concentrated by ethanol precipitation before sequencing.

DNA library preparation and sequencing

100 µl aliquots of purified DNA from all 60 samples were delivered to the Genomics and Microarray Core at University of Colorado Denver for library preparation and sequencing (Aurora, CO). Genomic libraries were prepared using the TruSeq DNA Library Preparation Kit (Illumina, Inc.). Next-generation sequencing was completed on the HiSeq 2000 (Illumina, Inc.) with 5 samples per lane, V4 chemistry, and paired-end reads of 125 bp in length.

Trimming and filtering of metagenomic sequence data

De-multiplexed sequence reads from libraries sequenced on the HiSeq 2000 were processed using the AMRPlusPlus bioinformatic pipeline (49). Briefly, the Trimmomatic software (122) was used to remove low quality sequences, and the "ILLUMINACLIP" command was employed to remove Illumina TruSeq adapters added during library preparation For each read, the first 3 and last 3 base pairs were removed. Then, starting at the 3' end of the read, a sliding window encompassing 4 nucleotides calculated if the average Phred score was lower than 15, in which case the 3' most nucleotide was removed and the window was moved forward until the average R score above 15. A Phred score is generated for each sequenced nucleotide by the Illumina HiSeq 2000 and a score of 15 signifies 96.8% nucleotide call accuracy. Based on Trimmomatic's default settings, any reads with fewer than 36 nucleotides were removed from further analysis. In order to remove any DNA that may have come from the host (i.e., bovine DNA), the remaining reads were aligned to the reference *Bos taurus* genome (123) using the

Burrows-Wheeler aligner (BWA) (93) with default settings for paired-end reads. Reads aligning to either of these genomes were removed from samples to create the "non-host reads" that were used for subsequent analysis. Sequencing results resulting from the number of raw, trimmed, and filtered reads and the average Phred score for each sample were compared using the generalized linear models with the "glm" function and the R platform (126) to assess systematic sequencing bias across sequencing batches. Similarly, differences in sequencing results between sample groups were tested with the Wilcoxon signed-rank test when comparing paired values from the same animal (arrival to day 11) and the Wilcoxon rank-sum test was employed when comparing animals at either time point.

Microbiome - Classification of bacterial sequences and identification of Salmonella enterica

Kraken 2 (128) was used to assign taxonomic labels to shotgun metagenomic DNA sequences using NCBI's reference genome database, RefSeq. RefSeq represents the most comprehensive, integrated, well-annotated set of genomes that includes viruses, archaea, and bacteria. Kraken uses the metagenomic sample reads that are typically 125 nucleotides each, partitions the reads into pieces that are 31 nucleotides in length, and searches for exact matches to the RefSeq reference database. Every match is scored with kraken's lowest-common-ancestor algorithm and the read is classified to the taxonomic level with the most points. The number of samples with reads classified as *Salmonella enterica* were identified and sample prevalence results were compared in contingency tables for shotgun metagenomics, aerobic culture and PCR. Reads classified as *S. enterica* were re-classified using the complete NCBI's nr/nt database using BLAST. Results suggested that plasmid sequences were being misclassified as *S. enterica* so to increase classification accuracy, we created a kraken "modified database" consisting of

curated bacteria, archaea, and viral genomes from RefSeq for a total of 5,200 genomes. Plasmid sequences were extracted from the genomic files and assigned to the "unidentified plasmid" ID (NCBI:txid45202) to increase the specificity of taxonomic read classification and account for the horizontal transfer of plasmids in microbial communities (see full script at https://github.com/colostatemeg/meglab-kraken-custom-db). Kraken's highest confidence value of "1" was selected to increase the alignment score threshold required for species-level classification and increase the accuracy of classification at higher taxonomic levels. Without the confidence flag, all 60 samples were incorrectly reported as *S. enterica*-positive. To further improve specificity, reads classified as *S. enterica* were re-classified with BLAST and NCBI's nr/nt database to confirm that sequences are truly unique to *S. enterica* when considering all available sequences on NCBI (see full script at https://github.com/EnriqueDoster/MEG-kraken-species-ID).

RESULTS

Study population and study design

This study compared aerobic culture, PCR, lateral flow immunoassays, and shotgun metagenomic sequencing for *Salmonella enterica* identification in fecal samples collected from feedlot cattle. These samples are from a published investigation on the effect of metaphylactic treatment with tulathromycin (one of the most commonly used antimicrobial drugs in American beef feedlots) on the resistome and microbiome of feedlot cattle (56).

Sample collection, culture, lateral flow immunoassay and PCR results

Aerobic culture and LFIs were used to test for the presence *of Salmonella enterica* on 376 fecal samples collected from study cattle. Use of culture yielded 4 positive samples and

agglutination tests revealed that of 3 isolates recovered from arrival processing samples 2 were serogroup C1 (including from the animal unsuccessfully sampled on day 11) and another serogroup K isolate, while the isolate recovered on day 11 was serogroup C1 (Supplemental Table 3.1). Regarding LFI assay, results suggest it might not be appropriate for evaluation of feces of feedlot cattle due to a high number of false-positive samples compared to the gold standard method, aerobic culture (Table 3.2). In these 60 samples, aerobic culture and PCR had 100% concordance, suggesting a 5% (3/60) overall prevalence for *Salmonella enterica* during the study period.

Sequencing results

Shotgun metagenomic sequencing generated 5.89 billion reads (2.95 billion paired reads) across 60 samples with an average of 98.20 million reads per sample (range 26.98 – 160.71 M [Supplemental Table 3.1]). The average Phred quality score of raw reads across all samples was 35.2 (range 34.54 – 35.82). Because of the high average Phred scores across samples, only 3.82% of reads were removed for low quality (minimum per sample = 2.21%, maximum = 6.36%). Of the remaining reads, 0.03% (1.8 M reads) were identified as bovine DNA and removed from subsequent analysis; two samples were nearly 20% bovine DNA and the other 58 samples ranged from 0.03% to 4.57%. Overall, there was no evidence of systematic bias in the sequencing effort for all samples.

Identification of Salmonella enterica using shotgun metagenomics

Following quality-based read trimming and removal of host genetic contamination, the kraken 2 software (128) was used to classify shotgun metagenomic reads taxonomically with the

National Center for Biotechnology Information's (NCBI) (166) Reference Sequence Database (RefSeq) (96). The kraken 2 flag, '--confidence' was used with the highest value of "1" to increase the score required to meet the threshold for species level classification. On average, 99.8% of the reads in each sample were unclassified (minimum 97.89%, maximum 99.91%). In all, more than 7.3 million reads were taxonomically classified with an average of 122,900 reads per sample. Using kraken 2 to analyze these data, *Salmonella enterica* was identified in 18.3% (11/60) of samples, compared to 5% prevalence using culture or PCR (Supplemental Table 3.2). However, through further examination of the RefSeq database structure, we noted that plasmid sequences, which can be actively transferred between bacteria, are commonly included within the reference genome files for each species. Kraken then incorrectly classifies these plasmids as being conserved (species-specific) to the organism that carried the plasmid when it was sequenced and submitted to NCBI. Therefore, we modified the creation of the kraken 2 database by separating the plasmid sequences included with RefSeq genomes and re-assigning them to a single taxon for all plasmid and synthetic vector sequences. Following re-classification of reads with the modified database, only 10% (6/60) of samples were *S. enterica*-positive, suggesting increased relative specificity compared to the standard database (Table 3.3). The number of misclassified reads as plasmids, on average, made up 16.6% of each sample's total reads classified using kraken (Figure 3.1). Finally, to further investigate the specificity of species level identification with the modified database, all sequences classified as *S. enterica* were isolated, and classification was confirmed using BLAST (167) version 2.8.1 and NCBI's largest database of genetic sequences, nucleotide collection (nr/nt). Out of 6 samples positive based on kraken 2, only 2 samples had sequences that were confirmed to be unique to *S. enterica,* but these results were not concordant with the positive culture and PCR results (Table 3.3). Remaining reads were

misclassified at the species level with greater than 99% sequence identity, but there was no evidence of misclassification above the family level, Enterobacteriaceae.

DISCUSSION

Results from this study demonstrate that metagenomic sequencing does not provide results comparable to culture and PCR for *Salmonella enterica* detection. Metagenomics provides access to information unavailable via culture and PCR alone, and has immense potential to fill knowledge gaps in microbial genetics and ecology. Metagenomics will inevitably fit a complementary role in pathogen detection and surveillance as sequencing costs decrease, reference databases improve, and bioinformatic analyses are streamlined (168). However, results from this study demonstrate that metagenomic approaches are reliant on proper use of bioinformatic methods, availability of extensive databases, and presence of uniquely-identifying genetic sequences within the taxonomic tree; or alternatively use of long-read sequencing technology for metagenomic samples. Until these requirements are available, it will be necessary to combine metagenomic results with traditional methods.

Aerobic culture and PCR are the most commonly used techniques for *Salmonella enterica* identification; in this sample set, these methods provided 100% concordant results with 5% positive samples (3/60). Low prevalence impaired the ability to make formal statistical comparisons, but results suggest that the LFI assay evaluated in this study is unsuitable for detection of *Salmonella* in the feces of feedlot cattle. Standard metagenomic analysis with default kraken parameters identified DNA sequences from *S. enterica* in 100% (60/60) of samples. However, increasing the threshold score required for species-level classification decreased the percentage of *S. enterica* positive samples to 18.3% (11/60) of samples and these results were 18.1% discordant with PCR/culture; these results could be misconstrued to signify

an overabundance of foodborne pathogens in the beef production system. Bioinformatic processing was critical to this result, including the choice of parameter settings and reference database, and we identified that plasmid sequences can cause false-positive *S. enterica* identification. Namely, removal of these plasmid sequences during database creation led to increased relative specificity with only 10% (6/60) positive samples and 10% discordant results with PCR/culture. Further, reference databases are inherently incomplete and the difficulty in identifying discriminatory regions between bacterial genomes is underscored by limitations in classification of short-read sequencing. Re-analyzing the reads classified as *S. enterica* using the most comprehensive sequence database, nr/nt, only 3.3% (2/60) of samples contained sequences unique to *S. enterica.* This decreased the estimated prevalence of *S. enterica* closer to results from aerobic culture and PCR, but the biological relevance of these results is still unknown. One of the samples with Salmonella-specific reads contained twelve such reads, while the other sample contained only one and neither sample was culture or PCR positive for *S. enterica*. The low number of matching reads could be a function of low sequencing depth, and future shotgun metagenomic studies will require deeper sequencing depth for species-level identification.

Achieving accurate and biologically-relevant results from metagenomic analysis poses a challenge and opportunity to the scientific community. As this cautionary tale demonstrates, scientists must intensely scrutinize results obtained from metagenomic data, including fulsome discussion of the full range of possible reasons for why the results may or may not be valid. This point is especially salient as the research, medical and regulatory communities continue to discuss application of shotgun metagenomics for purposes of disease diagnosis and pathogen detection across a variety of environments. The possibilities of metagenomic data must not be allowed to overshadow the methodical yet critically important requirements of the scientific

approach. As this work demonstrates, our ability to merge highly innovative methods with practical applications will depend on a successful cooperation between scientists studying bacteria with traditional methods, those experimenting with a metagenomics approach, and scientists developing bioinformatic tools.

Table 3.1. Products administered to study cattle at the time of arrival-processing (Day 1).

Product type	Commercial name	Manufacturer	Volume per animal	Additional information
Antimicrobial*	Draxxin	Zoetis	8 cc	Macrolide antimicrobial for treatment of cattle at high risk for bovine respiratory disease (BRD).
Anthelmintic	Noromectin	Norbrooks Labs	7 cc	Ivermectin parasiticide for the treatment and control of internal and external parasites of cattle.
Anthelmintic	Safeguard	Merck Animal Health	18 cc	For use in beef cattle for the removal and control of lung, stomach and intestine worms.
Vaccine	BoviAnthelmintic-Shield GOLD	Zoetis	2 cc	Protects cattle from infectious bovine rhinotracheitis (IBR) and bovine viral diarrhea (BVD).
Vaccine	Vision® 7	Merck Animal Health	2 cc	For use in healthy cattle as an aid in the preventing disease caused by Clostridium spp.
Steroid implant	Revalor-XS	Merck Animal Health	Implant	Trenbolone acetate and estradiol. It improves rate of gain and feed efficiency.

*Only the treated group received the antimicrobial treatment

Table 3.2. Aerobic culture and lateral flow immunoassay results for *S. enterica* identification in all 376 fecal samples collected at arrival and day 11.

	Aerobic culture		
Lateral Flow	Positive	Negative	Total
Positive	0	63	63
Negative	4	309	313
Total	4	372	376

Table 3.3. Aerobic culture and PCR (100% agreement) compared to shotgun metagenomic analysis with Kraken for *S. enterica* identification in 60 samples. Results are compared between using the standard kraken database, a modified Kraken database which considers plasmid sequences as unique taxa, and the modified database with additional confirmation using BLAST and NCBI's nr/nt database. Kraken was run with the "--confidence" flag at the highest value, "1".

	Culture and PCR		
Kraken – standard db	Positive	Negative	Total
Positive	1	10	11
Negative	2	47	49
Total	3	57	60
Kraken – modified db			
Positive	1	5	6
Negative	2	52	54
Total	3	57	60
Kraken – modified db + blast confirmation			
Positive	0	2	2
Negative	3	55	58
Total	3	57	60

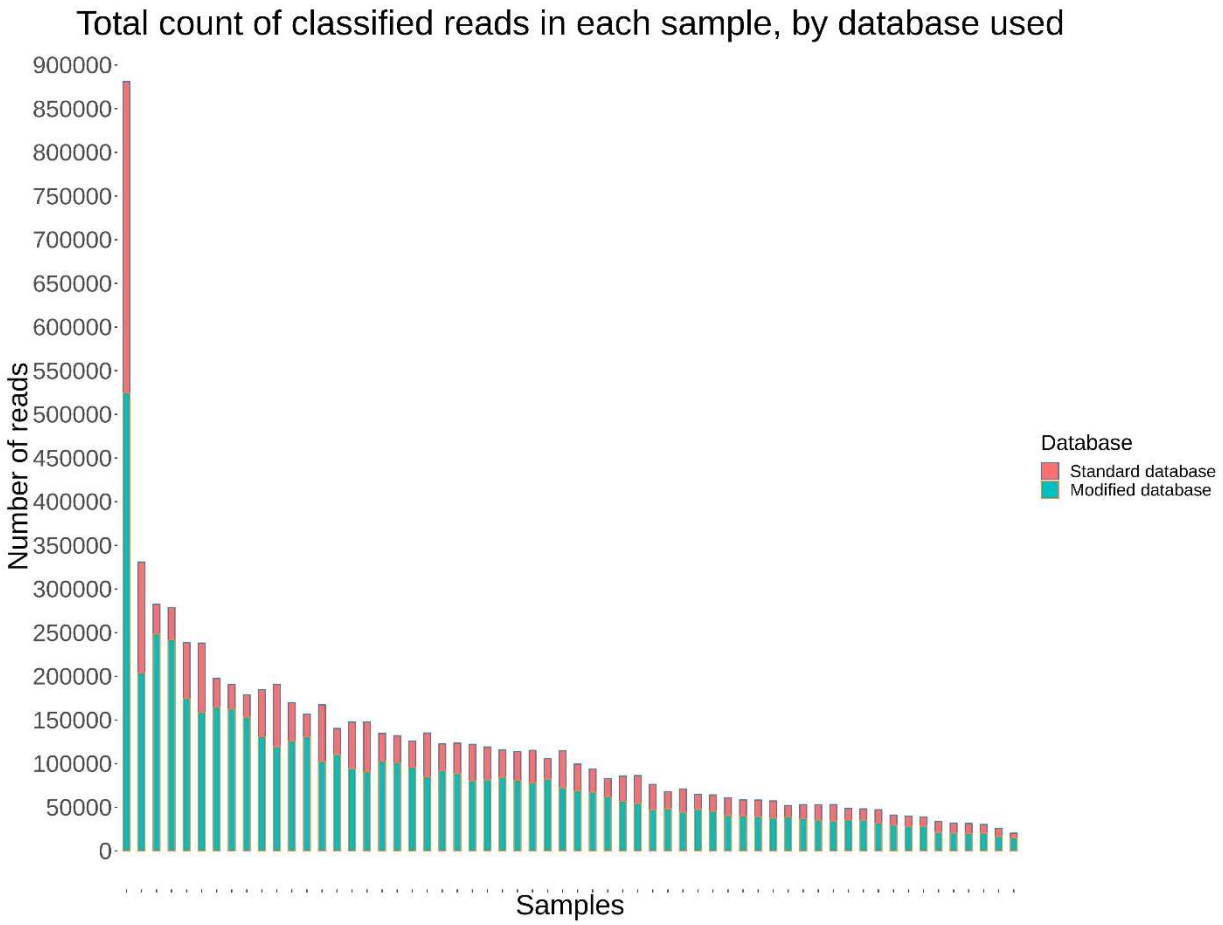


Figure 3.1. 60 individual samples on the x-axis with the total number of reads classified taxonomically using kraken in the y-axis. Reads classified using the standard database are shown in red and the decreased number of reads classified using the modified database are shown in green.

CHAPTER 4: ANTIMICROBIAL USE IN BEEF FEEDLOTS; EFFECTS ON THE MICROBIOME AND RESISTOME DYNAMICS IN INDIVIDUAL CATTLE

INTRODUCTION

Antimicrobial resistance (AMR) in bacteria is an ancient phenomenon that has emerged as a global public health threat due to widespread antimicrobial drug (AMD) use. Modern beef feedlot production systems administer AMDs to treat, control, and prevent disease, but studies raise the concern that AMD use may drive an increase in AMR bacteria that could be transmitted to people (16, 169) and cause treatment failure. Of particular interest, drug classes considered critically important for human medicine are used in beef production and could be selecting for bacteria resistant to these drugs (12). Globally, countries are attempting different methods to promote judicious use of AMDs, particularly to reduce the use for growth promotion in the US, Canada, Denmark, Norway and Sweden (17, 23, 24). Therefore, it is critical that we understand how AMD use in beef feedlot production affects AMR dynamics if we hope to ameliorate the risks of AMR promotion and transmission from livestock production systems.

Traditionally, previous AMR studies relied on aerobic culture to isolate a few bacterial species and test their resistance patterns with phenotypic testing (i.e. broth microdilution, Kirby-Bauer disk diffusion). Results from these tests provide details into AMR patterns for certain bacteria, but often lack the holistic perspective of considering the surrounding microbial community (microbiome) known to share a profile of AMR genes shared between species in this community (resistome). However, advances in high-throughput genetic sequencing allow us to characterize the genetic material in a sample and expand our understanding of how bacterial communities respond to AMD exposure in beef feedlot production systems.

Our study leveraged previously described fecal samples from a longitudinal study in four Canadian feedlots with detailed AMD use records (34). The gastrointestinal commensal species, *Escherichia coli,* and an upper respiratory pathogen, *Mannheimia haemolytica,* were isolated from rectal fecal samples and nasopharyngeal swabs, respectively. All isolates were tested for AMR using broth microdilution and Kirby-Bauer disk diffusion. The goal of this study was to use high-throughput genetic sequencing to investigate the effects of AMD use on AMR dynamics in the microbiome and resistome. Additionally, we compare our results with those gleaned from traditional methods to better inform future studies using these techniques to investigate AMR.

METHODS

Study Overview

Metagenomic sequencing was used to characterize the effect of AMD exposure on the fecal microbiome and resistome collected during a previously published 3-year longitudinal study of Canadian beef feedlot operations. Out of all pens of cattle arriving to the feedlot, 30% were randomly selected for inclusion into the study along with 10% of the animals in each pen. Fecal samples were collected per-rectum during arrival-processing at the feedlot (Arrival samples) and at a second time point > 60 days on feed (DOF) when all cattle were individually re-handled to conduct routine management procedures. All exposure to antimicrobial drugs, including parenteral treatments and in-feed exposures, was recorded and standardized across different drug classes using animal defined daily dose (ADD) (170, 171). Samples were processed for 16S rRNA amplicon sequencing and target-enriched shotgun to characterize the microbiome and resistome, respectively.

Sample Collection

As previously described (170), fecal samples were collected per rectum from individual cattle that were housed at four participating feedlots in western Canada. Cattle feces were collected when animals arrived at the feedlot, and again when they were re-handled at the facility. Samples were collected longitudinally from 8 October 2009 to 6 April 2010. Prior to shipment to the laboratory for processing, due to sample collection and shipment logistics, samples were kept refrigerated and preserved in in Cary-Blair transport medium as previously described (34) (https://bmcvetres.biomedcentral.com/articles/10.1186/1746-6148-9-216). Samples were archived in Whirl-Pak bags or 30mL medium transport tubes at -80°C until shipment on dry ice to Colorado State University (CSU) in May of 2017. At CSU they were stored again at -80°C until samples were individually thawed for DNA isolation and extraction. Each arrival sample and re-handling sample were labelled and archived together, so that these corresponding samples from each animal could be correlated after extraction.

Sample selection for DNA Isolation and metagenomic sequencing

Samples from a total of 60 animals were selected for use in this study, including fecal samples collected from two time points that were used for metagenomic sequencing. Samples were selected with stratified-random sampling primarily to capture different levels of parenteral ADD exposure (range: 0 - 7); no parenteral treatment, low parenteral exposure (<4 ADD), and high exposure (>4 ADD).

Fecal samples were thawed at 4°C and 5g was aliquoted into 50mL conical tubes. DNA was extracted the Qiagen PowerMax Soil Kit (Qiagen Laboratories) according to manufacturer's instructions, except for the first centrifugation. We found that samples provided a higher yield when centrifuged for 5 minutes in the PowerBead tubes, rather than the recommended 3 minutes.

Isolated DNA samples were eluted using 3mL of the kit's provided buffer solution and passed through the silica DNA filter twice for optimal yield. Following extractions, 6.6mL of molecular grade ethanol was added to each sample, as well as 300µL 0.3M sodium acetate (NaOAc), to facilitate DNA precipitation by ethanol. Precipitation was further facilitated by storing samples at -20°C overnight. Samples were then aliquoted into 2mL tubes, and centrifuged at 13,000 x g for 20 minutes, allowing a pellet of crude DNA to form. The supernatant was removed after each centrifugation, and 2000µL of each sample was added to the same tube with the DNA pellet and centrifuged again. This was repeated until the entire sample was centrifuged and the DNA allowed to form a single pellet for each sample. Once the pellet of crude DNA was formed, 700µL of 70% ethanol was added to wash the pellet by removing salts and centrifuged again at 13,000 xg for 10 minutes. The remaining supernatant was removed, the pellet was air-dried to insure no ethanol remained, and the DNA was resuspended in 150µL of the provided elution buffer.

Following resuspension, each extraction's concentration was quantified using the Qubit 2.0 Fluorometer and dsDNA High Sensitivity Buffer and Reagent (Thermo Fischer Scientific), according to the manufacturer's specifications. Concentration for each sample were averaged between two separate measurements. The concentration results were and sample purity (260/280 measurement) were verified with the use of the NanoDrop 1000 Spectrophotometer (Thermo Fischer Scientific). Average DNA yield was then calculated in micrograms. If a sample failed to reach a total DNA yield of >9ug, it was re-extracted. Because some re-extractions required more sample material than was provided, 2-5mL of PBS was used to rinse the Whirl-Pak bags or transport tubes each deficient sample was stored in. The volume of rinsate used was dependent upon the weight of the remaining sample, with more PBS used for samples with less weight, in

order to reach a 5g extraction volume. Pooled samples (30µL) were shipped on ice to the Novogene Corporation for 16S sequencing.

Library Preparation and Metagenomic sequencing

For microbiome sequencing, we sent between 200-500 ng of DNA from each sample (n=120) to the Novogene Corporation for 16S rRNA gene amplification and sequencing. The V4 region of the 16S subunit was amplified with the primer set 515F/806R [5'-GTGCCAGCMGCCGCGGTAA-3']/[5'-GGACTACHVGGGTWTCTAAT-3'], with a fragment length of 292 bp. Amplicon sequencing was performed on the Illumina HiSeq 2500 Sequencing System to produce paired end 250 bp reads (PE 250) at a targeted read depth of 100,000 PE reads per sample.

SureSelectXT HS Reagent Kit for Illumina Paired-End Multiplexed Sequencing Library (Agilent Technologies) was used to prepare samples for target-enriched resistome sequencing. A custom bait design targeting AMR genes, 'MEGaRICH' (84) was used to improve "on-target" sequencing and ameliorate the challenge of sequencing microbes from a sample primarily consisting of host DNA. Denver Genomics and Microarray Core Facility (Denver, CO) sequenced all study samples using the NovaSeq 6000 Sequencing System (Illumina), targeting a read depth of 100 million paired end reads per sample of 150bp in length.

Analysis of sequence quality

Summary statistics regarding the number of raw, trimmed, and non-host reads for each sample were compared using generalized linear models with the "glm" function and the R platform (126) to assess systematic bias across the following sequencing metadata: sequencing

run, batch, and lane. For study design metadata, primary comparisons of interest were between varying antimicrobial drug exposure levels and DOF when the sample was collected (sampling time). Differences in sequencing results between sample groups were tested with the Wilcoxon signed-rank test when comparing paired values from the same animal and the Wilcoxon rank-sum test was employed when comparing between sampling time.

Microbiome and Resistome characterization

To describe the profile of microbes present in feedlot cattle feces, reads from 16S rRNA amplicon sequencing were analyzed using the collection of tools contained in Quantitative Insights Into Microbial Ecology version 2 (Qiime2-2017.12) (172). Briefly, all reads were processed for sequence quality and denoising using DADA2 (Callahan et al., 2016), taxonomic classification using a naive bayes classifier trained on the GreenGenes database (97), and removal of chloroplast and mitochondrial DNA contaminants. Results were then exported into count tables and summarized using the R statistical software.

To identify the resistome in feedlot cattle feces, the targeted AMR metagenomic sequencing samples were processed using the AMRPlusPlus bioinformatic pipeline and the MEGARes resistance database (49). Two updates were added to AMRPlusPlus pipeline; one to deal with PCR duplication introduced with targeted AMR amplification and another to improve the classification of AMR genes requiring the presence of specific SNPs. Further details on the pipeline can be found in the documentation website (http://megares.meglab.org/AMRPlusPlus). Read trimming and quality filtering was performed using Trimmomatic (122). Host contamination was identified using the Burrows-Wheeler Aligner (BWA) software (93) with alignment to the Bos Taurus genome (123) and removal of those reads with SamTools (124).

These non-host reads were then aligned to the MEGARes database with BWA. Read de-duplication was performed with SamTools on the resulting .sam files from alignment to MEGARes.

Only genes accessions with reads aligning to >80% of the reference nucleotide sequence were considered for further analysis, with the exception of reads aligned to genes that require the presence of specific single nucleotide polymorphisms to confer resistance. These reads were identified, pulled from the samples, and re-classified separately using Resistance Gene Identifier (95) and the "strict" paradigm which incorporates detection models and CARD's curated similarity cut-offs to increase accuracy in identifying functional AMR genes (see full script: https://github.com/EnriqueDoster/MEG-AMRPlusPlus-RGI-SNPconfirmation). A list of important AMR gene determinants in human-associated pathogens were identified a priori and searched for in all samples: (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat, cfr).

Count matrix processing

Differences in the microbiome and resistome of individual feedlot cattle were analyzed using the R statistical programming software. Cumulative sum scaling (CSS) (131) was used to normalize counts and account for differences in sequencing depth. The resistance data was then summarized to the class and mechanisms level to avoid bias at the "gene" level associated with irregular naming criteria for new resistance genes (132). Statistical analysis was accomplished using the R packages "metagenomeSeq" and "vegan" (130, 131). Sparsely represented resistance gene accessions which were identified in fewer than 5% of samples were removed from further analysis based on published recommendations (131). The taxonomic lineage was identified for

each taxon in the microbiome and counts were aggregated to these 6 Linnaean taxonomic levels: phylum, class, order, family, genus, and species. In total, there were 6 count matrices for the microbiome, but to reduce the repetitive reporting of results at all levels and because results at lower taxonomic levels are not considered very reliable (133), statistical results for microbiome are presented at the phylum, class and order levels. Richness and Shannon's diversity indices were calculated for each sample using "vegan" at all levels. In total, 8 unique normalized count matrices (i.e., 6 count matrices describing the microbiome and 2 count matrices characterizing the resistome) were analyzed and reported.

Statistical analysis

The primary analysis of interest was between sampling time (Arrival vs Re-handling) and total ADD exposure categories. The total ADD exposure for each animal was calculated as the sum of ADDs from all sources and categorized into 3 categories based on ADD sample distribution; Low ADD exposure (< 8), medium ADD exposure (8 – 18, and high ADD exposure (> 18). Similarly, the DOF for each animal at the time of sample collection during Re-handling was categorized into 5 ranges (arrival - 3 DOF, 4-70 DOF, 71-120 DOF, 121-180 DOF, and > 180 DOF). Diversity indices were statistically compared using the Wilcoxon signed-rank test ("wilcox.test" function in R) for samples from the same animal and "glm" to test differences between other sample groups. CSS- normalized counts were Hellinger-transformed (135) for ordination using the metaMDS function from "vegan", which employs non-metric multidimensional scaling on Euclidian distances. Analysis of similarities (ANOSIM) (136) was used to test differences in the microbiome and resistome between categorical metadata sample groups (e.g., Arrival vs Re-handling samples, ADD exposure category, and DOF sampling

category). MetagenomeSeq's "fitZig" function was used to fit a zero-inflated Gaussian model and compare log2-fold differences (131) in microbiome and resistome features between sampling time. Limma's "makeContrast" function (137) were then used for pairwise comparisons, P-values were adjusted for multiple tests using the Benjamini-Hochberg procedure (138), and alpha=0.05 was selected as the statistical significance cut-off value. To account for spurious statistically significant differences in low abundance features, only features with an average expression > 1 were considered.

Raw counts were hellinger-transformed (135) and redundancy analysis was performed on the microbiome and resistome composition to further evaluate the potential significance of different AMD use practices using the "rda" function in R. Significance of the correlation between independent variables and the variance in the microbiome and resistome composition were then tested using the "anova" function in R. With the goal of characterizing the effect ADD exposure and time in the feedlot (days on feed or DOF) have on the microbiome and resistome, samples were grouped into 18 metadata categories for analysis. Values for ADD exposures were aggregated by route of administration (in-feed vs parenteral) and by drug class including macrolides-lincosamides-streptogramin (MLS), tetracyclines, phenicol, and bactrim (sulfamethoxazole and trimethoprim combination). Samples were summarized into metadata variables that reflect the amount and type of antimicrobial drug exposure as well as time in the feedlot and days since the most recent parenteral treatment (Supplemental Table 4.1). Variables included:

"Sampling_Time","Total_ADD","Feed_ADD","Parenteral_ADD","DOF","num_tx","Days_since_tx","total_tetracycline_ADD","total_MLS_ADD","feed_MLS_ADD","feed_tetracycline_ADD",

"parenteral_tetracycline_ADD","parenteral_MLS_ADD"parenteral_phenicol_ADD","parenteral_sulfonamide_ADD", and “Feedlot_ID”. All of these variables were included in the starting model for step-wise variable selection and anova testing to identify the best fitting model. Additionally, an *a-priori* model was included to test the effect of time (DOF), in-feed ADD exposure and parenteral ADD exposure.

RESULTS

Sequencing results

In this study we employ metagenomic sequencing to characterize the effect of AMD use on the microbiome and resistome in feces collected during a previously published 3-year longitudinal study of Canadian beef feedlot operations. Out of 120 samples selected for inclusion in this study, only 94 were successfully sequenced using AMR targeted-enrichment (resistome) because of low starting DNA concentrations not meeting the threshold for AMR targeted-enrichment. All 120 were successfully sequenced using 16 rRNA amplicon sequencing (microbiome). Across the remaining 94 samples, >1.49 billion paired end reads were produced (mean: 15,926,612, range: 3,113,837 – 25,239,487 [Supplemental Table 4.2]). Filtering to improve overall read quality and to remove bovine host DNA a total of 336,463,224 reads were excluded and on average, these reads made up 22.9% of each sample (range: 3% - 38.5%). There was a statistically higher number of raw sequencing reads in samples from the second time point (mean = 16,997,789) compared to arrival samples (mean = 14,855,436). There were no significant differences in sequencing the microbiome. More than 17.6 million reads were produced from 16S rRNA sequencing across 120 samples (mean: 147,046, range: 101,543 – 208,020). Quality filtering, identification of sequence variants with DADA, and removal of

chloroplast and chimeric sequences identified a total 4,902,718 sequence variants with each sample averaging 40,855 unique sequence variants per sample (range: 20,943 – 65,270).

Resistome results

Following alignment of reads to the MEGARes AMR database, removal of duplicate reads, and re-classification of reads aligning to gene accessions requiring SNP confirmation with RGI, a total of 3,548,954 alignments to AMR gene accessions were identified across 94 samples (mean: 3,548,954 reads per sample, range: 11,635 – 129,357). These counts were classified as representing 1,608 different gene accessions in the MEGARes database, confering resistance to 26 different drug classes through 80 distinct resistance mechanisms. Out of the 26 drug classes identified, half were in low abundance across all samples and together made up 0.05% of reads aligning to AMR gene accessions. Regardless of the time of sample collection, the seven most abundant drug classes were tetracyclines (46%), multidrug resistance mechanisms (such as multidrug efflux pumps - 16.7%), aminoglycosides (12.5%), macrolide-lincosamide-streptrogramin (MLS) (6.6%), sulfonamides (6.4%), betalactams (6.3%), and phenicol (4.1%) with the remaining 19 classes each making up less than 1% of classified reads. Of the alignments to genes that confer tetracycline resistance, 69.6% of all alignments were to tetracycline resistance ribosomal protection proteins and 27.5% were to major facilitator superfamily (MFS) efflux pumps. In the second most abundant group of resistance determinants, multidrug resistance, 65.4% of alignments were to the multidrug efflux pumps and 34.6% to MDR regulator mechanisms.

While tetracycline and multidrug mechanisms dominated the resistome composition in all samples, multidrug resistance mechanisms were less abundant when cattle were resampled at the

second time point such that aminoglycoside resistance became the second most abundant drug class (Figure 4.1). ANOSIM testing suggests that the overall resistome composition shifted significantly over time at the class level (ANOSIM R = 0.32, P = 0.001) and mechanism level (ANOSIM R = 0.34, P = 0.001) (Figure 4.2). Despite these changes in resistome composition, richness and Shannon's diversity indices did not undergo a significant change between sampling time. The ZIG model identified that out of 15 drug classes with an average expression > 1, the relative abundance of 12 drug classes shifted significantly between sampling times. Most commonly, these shifts were related to increased abundance of alignments to tetracyclines, MLS, sulfonamides and phenicol resistance alignments increased over time and a decreased abundance for hits for 8 other drug classes (P-value < 0.05).

To further investigate how much of the variation in the resistome composition among different samples could be explained by independent variables (sampling time, feedlot ID, AMD exposure in ADD units, etc.), redundancy analysis was performed starting with a full model containing all 18 metadata categories (Supplemental Table 4.1). Step-wise selection to identify a model that maximized the amount of explainable variation identified sampling time as the only significant variable (ANOVA P < 0.05); this model explained 2.5% of the constrained variance and 8.6% of the total variance in resistome composition (Figure 4.3). Because of the significant shifts in resistome composition over time, the resistome of animals collected at the second time point were similarly analyzed with redundancy analysis. This best-fitting redundancy model for samples collected at the second time point included only two variables (ANOVA P < 0.05): DOF and in-feed MLS ADD exposure. The final model explained 4.6% of the variation in the resistome and suggests that only 0.49% of that variation can be explained by the independent variables measured in this study. Of the constrained variance, DOF and in-feed MLS ADDs

explain an equal amount of variance (~0.25) though neither tested statistically significant ($P < 0.07$). In-feed MLS ADDs were only given to 11/47 animals in the study compared to 30/47 receiving parenteral treatment with a macrolide AMD and all 47 animals exposed to in-feed tetracyclines.

Of the *a priori* list of 13 genes identified as important to medicine when expressed in human pathogens, all genes were identified in at least one sample in our study (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat, cfr). Across 47 samples, 204,934 reads mapped to these genes and were predominantly (>98%) represented by alignments to cfr, bla(OXA), bla(CTX-M), bla(TEM), and vat. Between sampling time points, the number of animals with alignments to these 5 genes generally increased over time and decreased for the remaining 8 genes investigated as being of high importance (Figure 4.4). Though samples were collected from animals located in 4 different feedlots, similar patterns emerged in the prevalence of these genes over time in the feedlot. The bla(TEM) and cfr group of gene accessions were in the highest relative abundance of all five genes and could be identified in 85.1% (40/47) and 95.7% (45/47) of arrival samples, respectively. Though the low abundance of normalized counts precluded the ability to statistically compare logFold changes in abundance over time, the sum of alignments decreased for bla(TEM) and cfr (FIG). Alternatively, prevalence and alignments to bla(OXA), bla(CTX-M) and vat gene accessions increased over time.

Microbiome results

A total of 794,953 reads were classified taxonomically into 33,623 amplicon sequence variants. Together, these taxa represented 29 phyla, 63 classes, and 100 orders. Three phyla,

Proteobacteria, Firmicutes and Bacteroidetes, predominated in the microbiome and accounted for > 93% of all normalized counts (45.4%, 36.9%, and 10.9%, respectively). These 3 phyla were each predominantly represented by a single taxonomic class which contained a majority of hits within these phyla. The Clostridia class made up 74.6% of all Firmicutes alignments and the Gammaproteobacteria class was 99.4% of all Proteobacteria across all samples. Pseudomonadales (44.7%), Clostridiales (26.9%), Bacteroidales (8.7%), Lactobacillales (6.6%), RF39 (2.7%) and Flavobacteriales (2%), and Enterobacteriales (1.1%) were the most abundant taxa at the order level, with the remaining taxa each making up less than 1% of classified reads.

Overall, the microbiome composition in feedlot cattle were dominated by Proteobacteria, Firmicutes, and Bacteroidetes at both sampling time points with the majority of the identified phyla accounting for < 0.01% of all classified reads (Figure 4.5). Though there were no significant differences in richness or Shannon's diversity, ANOSIM testing confirms that microbiome composition shifted significantly between sampling time points at the phylum (ANOSIM R = 0.19, P < 0.01), class (ANOSIM R= 0.21, P < 0.01), and order (ANOSIM R= 0.22, P < 0.01) taxonomic levels (Figure 4.6). Out of the 10 phyla with an average expression >1 as calculated by the ZIG, 7 phyla significantly changed in relative abundance between sampling times (P-value < 0.05). Bacteroidetes, Proteobacteria, and Spirochaetes significantly increased in relative abundance between the first and second sampling, with an accompanying decrease in Firmicutes, Cyanobacteria, Actinobacteria, and Verrucomicrobia.

Redundancy analysis and step-wise selection of exposure variables that best explain the variance in the microbiome resulted in a statistically significant model (ANOVA P < 0.05) explaining 1.7% of the variance between all samples (Figure 4.7). Of the 4 variables included in the model, only sampling time and in-feed ADD were significantly (ANOVA P < 0.05)

associated with explaining 1.2% and 0.2% of the microbiome variance, respectively. By the second sampling time point, however, 9.6% of the total variance in the microbiome could be explained with a redundancy model containing two variables, in-feed tetracycline resistance ADD and total in-feed ADD. Only total in-feed ADD was significantly associated with representing 1% of the constrained variance.

DISCUSSION

Our results suggest that, as evaluated by metagenomic sequencing, exposure to antimicrobial drugs might exert a greater effect on the microbiome than on resistome composition in beef feedlot cattle. However, this effect is likely small compared to impacts associated with the amount of time cattle spent in the feedlot as a marker for other un-measured factors in the environment that could impose a greater overall impact. The largest measurable association was sampling time (Arrival vs Re-handling) which accounted for 2.5% of the variation in the resistome, but these changes in the resistome were not associated with increasing DOF or AMD exposure at the second sampling point. Major shifts that are not related to AMD use that could account for the differences between arrival and re-handling resistome composition may be related to a multitude of environmental pressures involved with initial transportation to the feedlot, processing at the feedlot and diet changes once in the feedlot (173). The microbiome likewise was also significantly influenced by sampling time, but additionally by the total amount of time cattle spent in the feedlot. Comparing only samples from the second sampling time, in-feed antimicrobial drug exposure explained only 1% of differences in the microbiome of samples.

It is important to note that our study faces the same limitations of many high-throughput sequencing studies. Sequencing depth could have been inadequate to identify more subtle

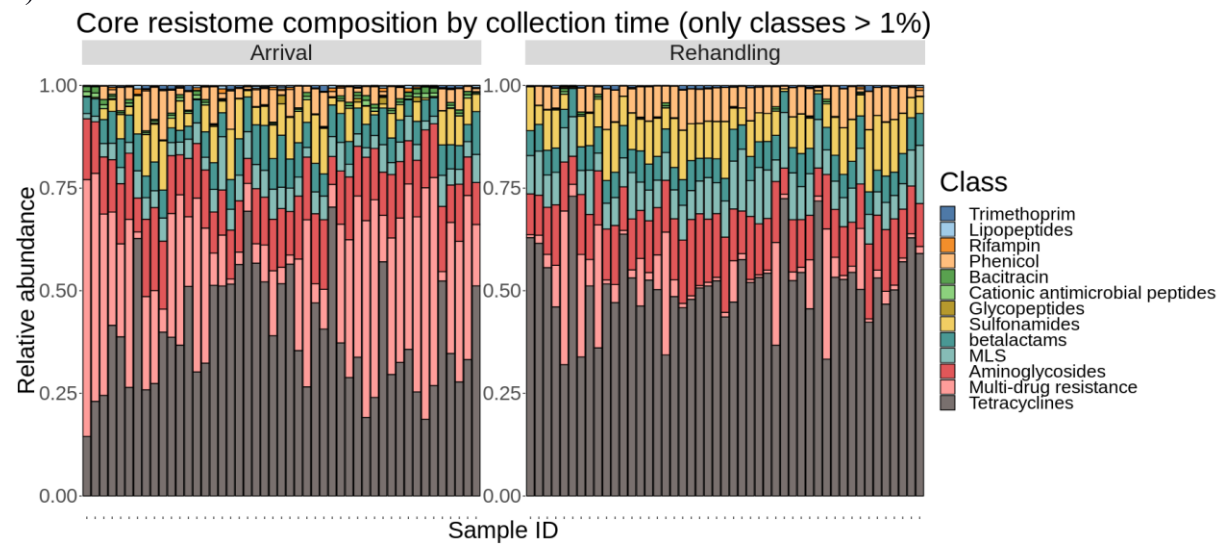
dynamics occurring in low abundance or “rare” features. Additionally, classification of AMR gene accessions in our samples does not ensure the functional activity of these genes and instead serves only to identify the potential resistome function. This could misrepresent instances in which gene expression of AMR determinants is not represented by genetic composition. Lastly, this study leverages previously collected samples and while this serves as an example of the benefit of using metagenomics on archived samples to add a different perspective it also impacts the potential external validity of microbiome and resistome results garnered from samples not originally processed for DNA extraction and sequencing (77). In our study, fecal samples were originally processed for aerobic culture, stored in Cary Blair media in a refrigerator prior to freezing. Therefore, our internal validity is sound as all samples were processed with the same protocols, but comparisons to external studies should be done with caution.

In parallel with characterizing broad patterns in resistome changes, high-throughput sequencing allows interrogation for the presence of specific genes with medical importance. Among the 13 genes considered *a priori* to be of high importance to public health, we identified all of these in the sample set of 94 samples, but five (cfr, bla(OXA), bla(CTX-M), bla(TEM), and vat) were identified in at least 32/47 (68%) samples at the second sampling time. The function of all these medically important AMR gene accessions cannot be inferred from the methods employed in this study and future studies will require incorporating multiple techniques to better characterize the risk of AMR emergence.

A thorough search of the current relevant literature yielded no other metagenomic studies of this kind in individual beef feedlot cattle. Paired with extensive antimicrobial drug use records across 4 different feedlots and 60 animals sampled at two time points, we characterize general trends observed and provide a snapshot into the dynamic microbial ecology involved in livestock

production. The power of metagenomic sequencing is likely to be best employed for the culture-independent surveillance of AMR genes and this study provides an overview of how the resistome and microbiome of feedlot cattle are influenced by antimicrobial drug use practices. With increasing understanding of how herd management decisions can influence the microbiome and the resistome, practices can be identified that maintain the balance of benefiting from the use of antimicrobial drugs without increasing the risk of AMR emergence in pathogens. As we learn to better manage AMR through livestock production practices, metagenomic analysis will be critical for incorporating a holistic perspective into community-wide changes.

A)



B)

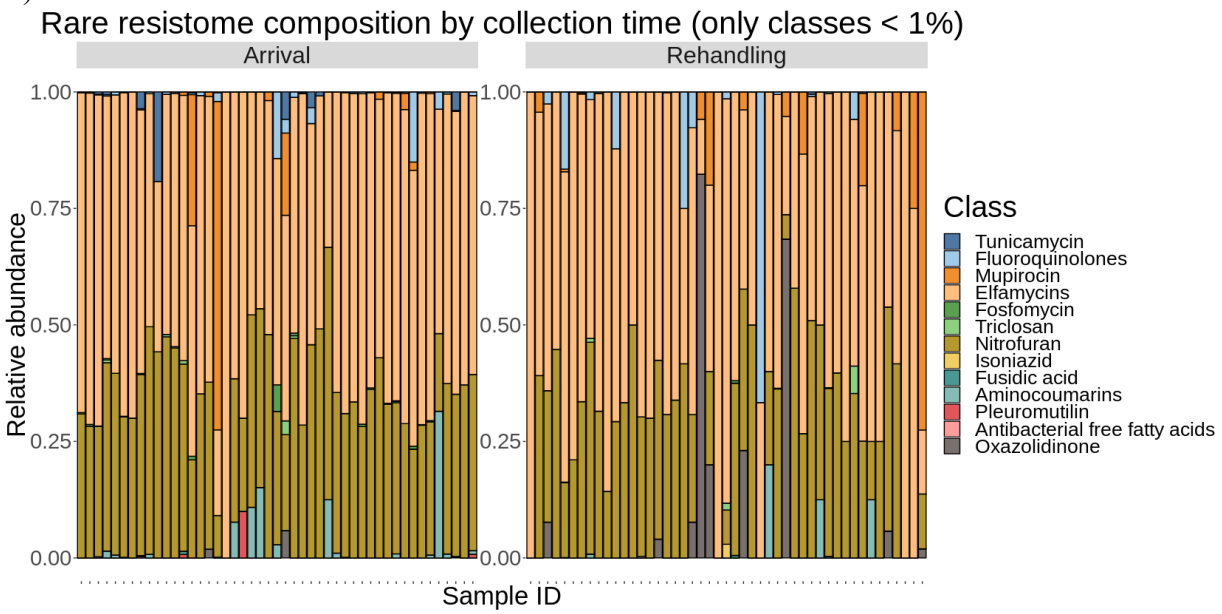


Figure 4.1. Resistome composition at the drug class level for all samples by sampling time point; Arrival at feedlot or later during the feeding period at Re-handling. A) Shows the 13 drug classes present in a relative abundance > 1% of all resistome alignments (Core) and B) is the relative abundance for the 13 drug classes in low relative abundance (Rare).

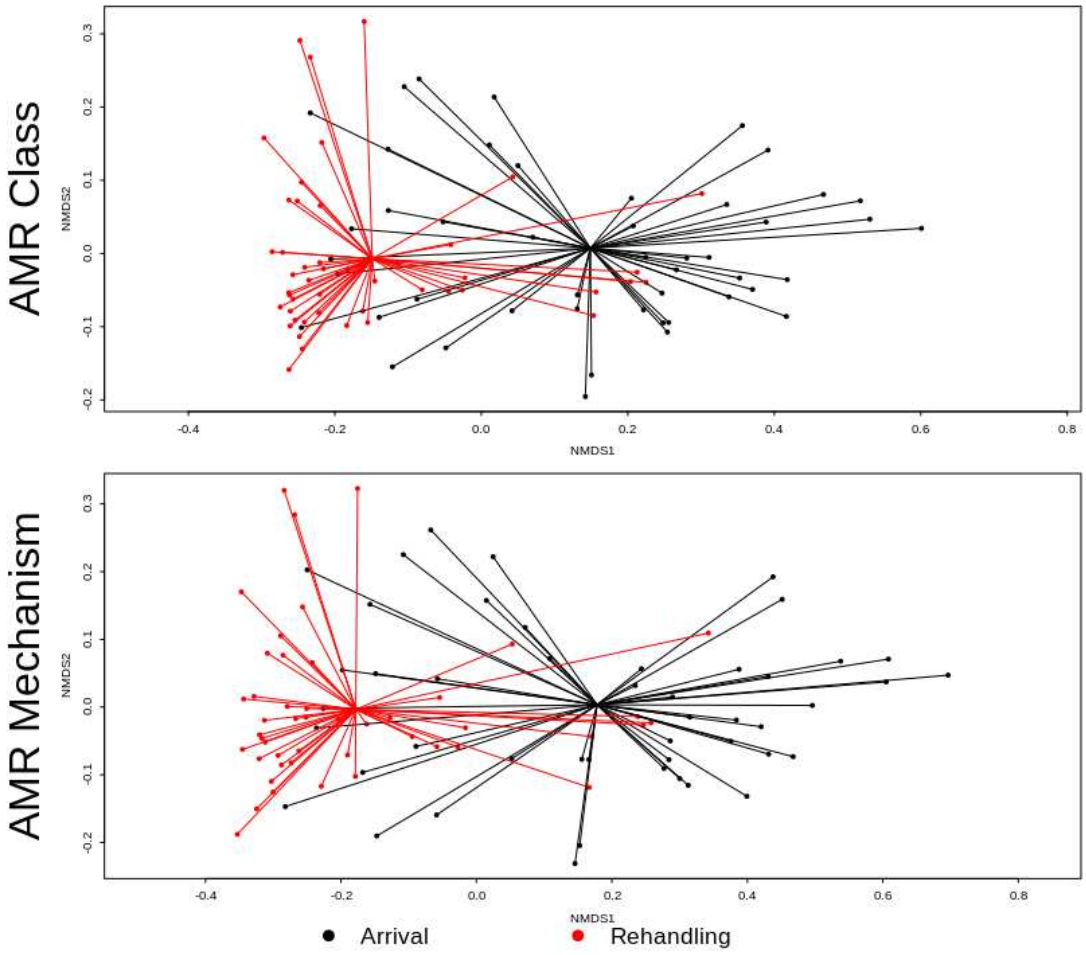


Figure 4.2. Ordination comparing resistome composition at the AMR drug class and resistance mechanism, using non-metric multidimensional scaling (NMDS), for the two study groups at Arrival and Re-handling samples. Separation of resistomes between sampling time was statistically significant at the class (ANOSIM R = 0.32, $P < 0.05$) and mechanism levels (ANOSIM R = 0.34, $P < 0.05$).

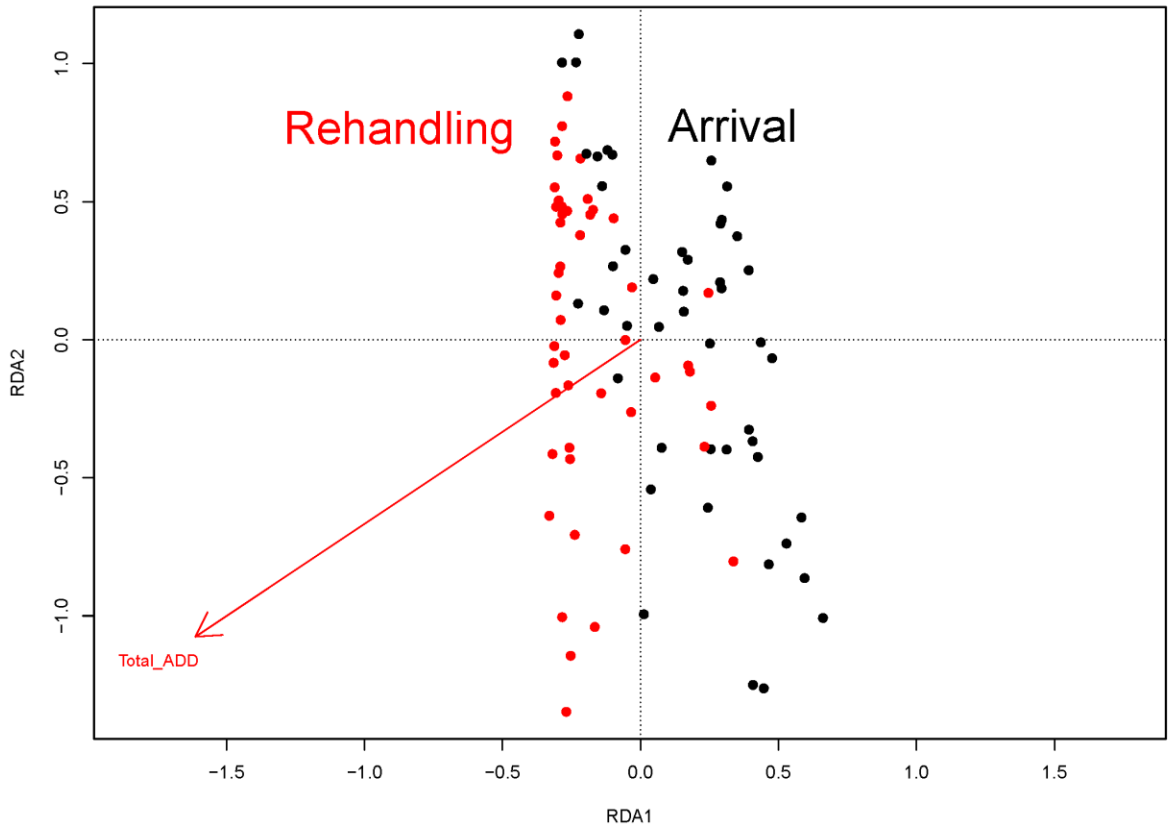
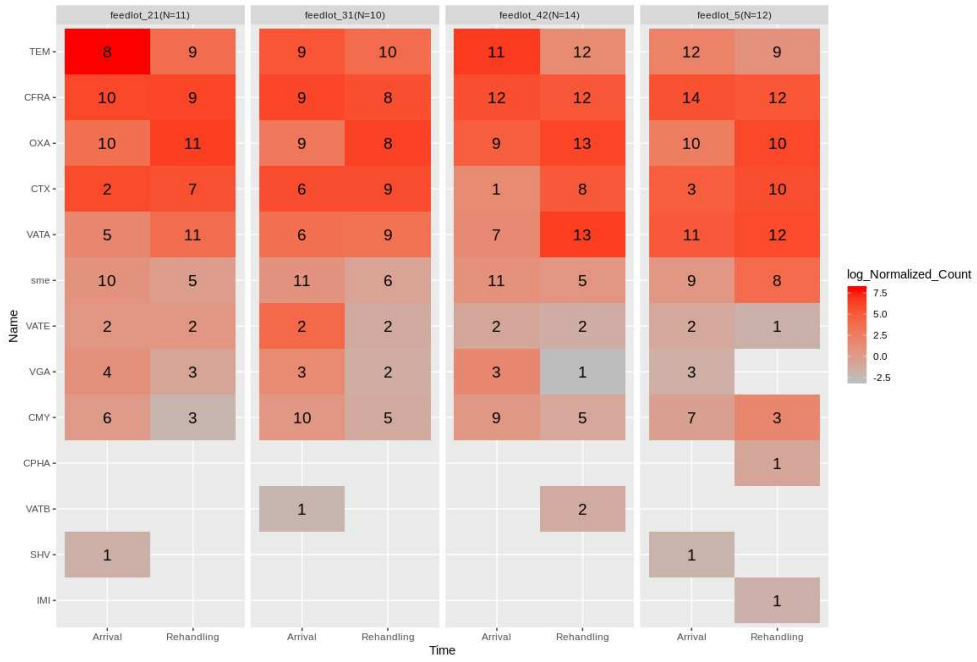


Figure 4.3. Redundancy analysis (RDA) plot illustrating the relationship between resistome composition and antimicrobial drug exposure metadata. Step wise model selection was used to determine the combination of variables that best explain the variance in the resistome. Sampling time and total exposure to antimicrobial drugs (in ADDs) were included in the final model but only sampling time was statistically associated with explaining 2.5% of the variance in the resistome across all 94 samples.

A)



B)

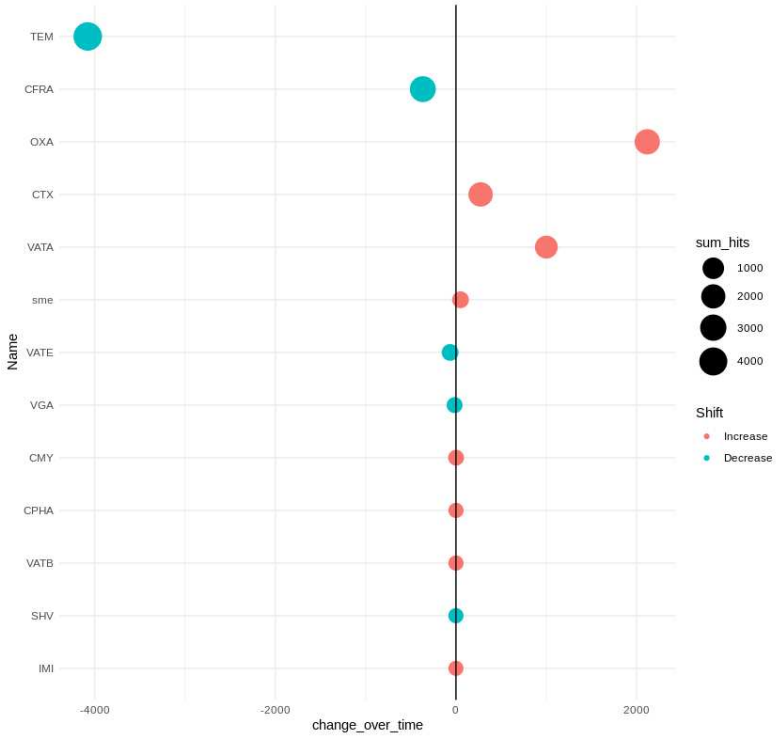
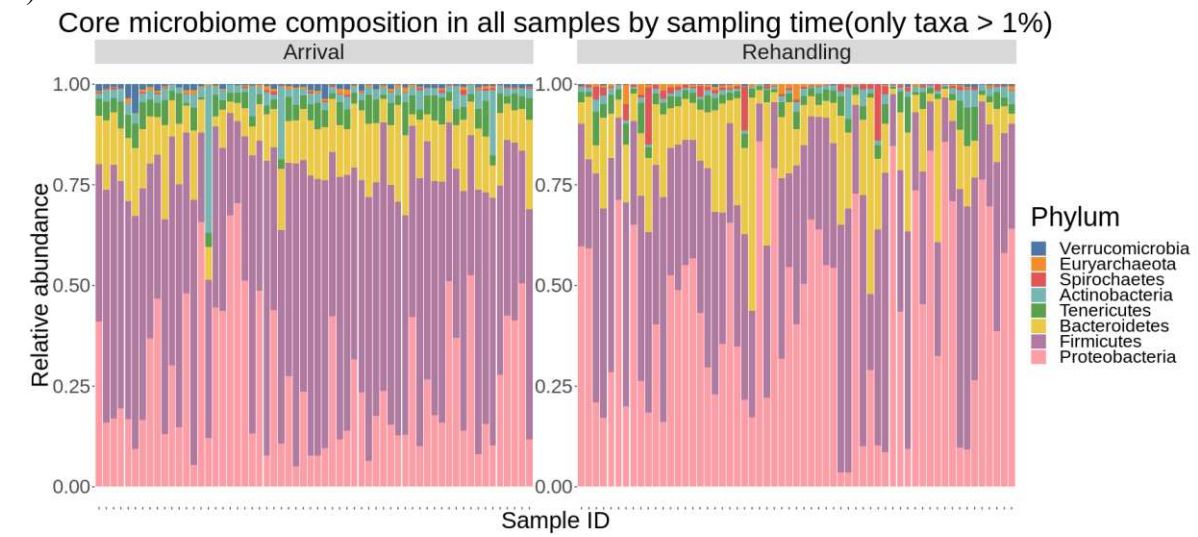


Figure 4.4. A) Heatmap representing the prevalence of medically important AMR genes across 4 different feedlots. The increasing shade of red represents the relative abundance of counts for each gene group on the rows. The number on each cell signifies how many samples contained alignment to that gene group. B) Plot of the change in alignments from the arrival samples to exit for each medically important AMR gene group across all samples.

A)



B)

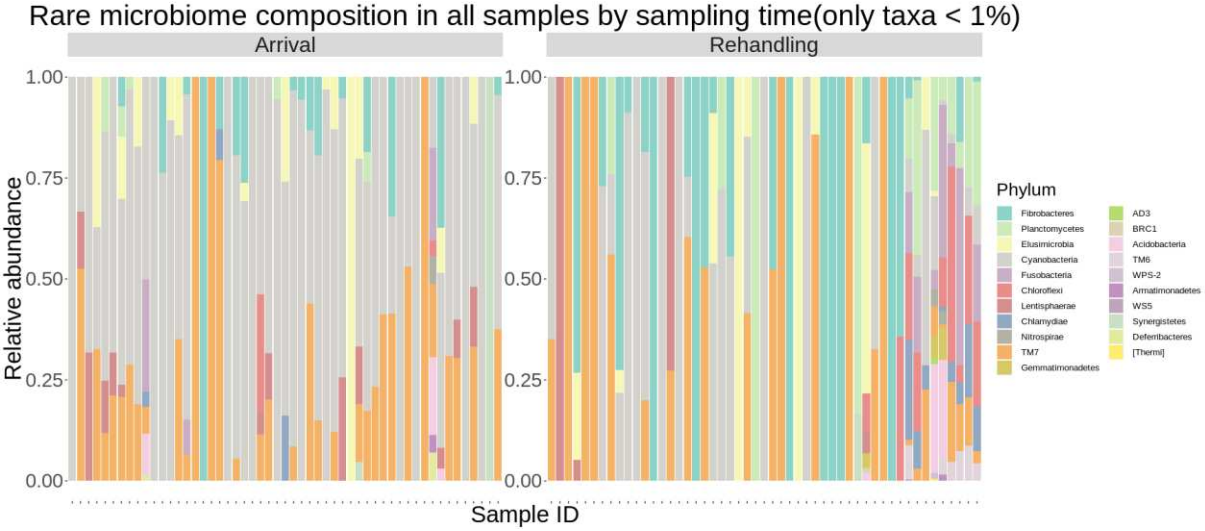


Figure 4.5. Microbiome composition at the phylum taxonomic level for all samples by sampling time point; arrival at feedlot or at a second time point closer to exit of the feeding period. A) Shows the 8 phyla present in a relative abundance > 1% of all microbiome alignments (Core) and B) is the subset of 21 phyla in low relative abundance (Rare).

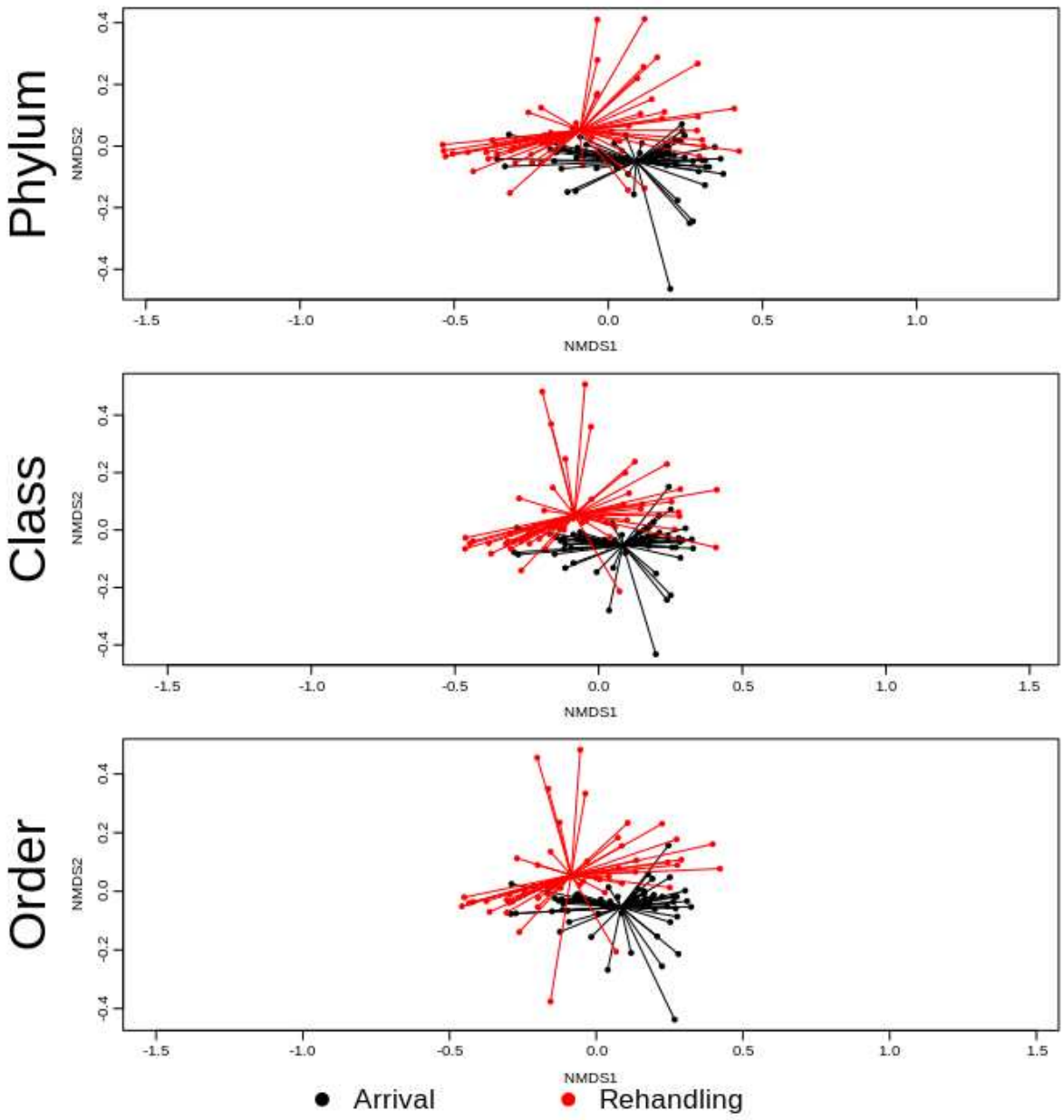


Figure 4.6. Ordination comparing microbiome composition at the phylum, class, and order levels using non-metric multidimensional scaling (NMDS), for samples collected at Arrival and Re-handling. Separation of resistomes between sampling time was statistically significant at the phylum (ANOSIM $R = 0.19$, $P < 0.01$), class (ANOSIM $R = 0.21$, $P < 0.01$), and order (ANOSIM $R = 0.22$, $P < 0.01$) taxonomic levels.

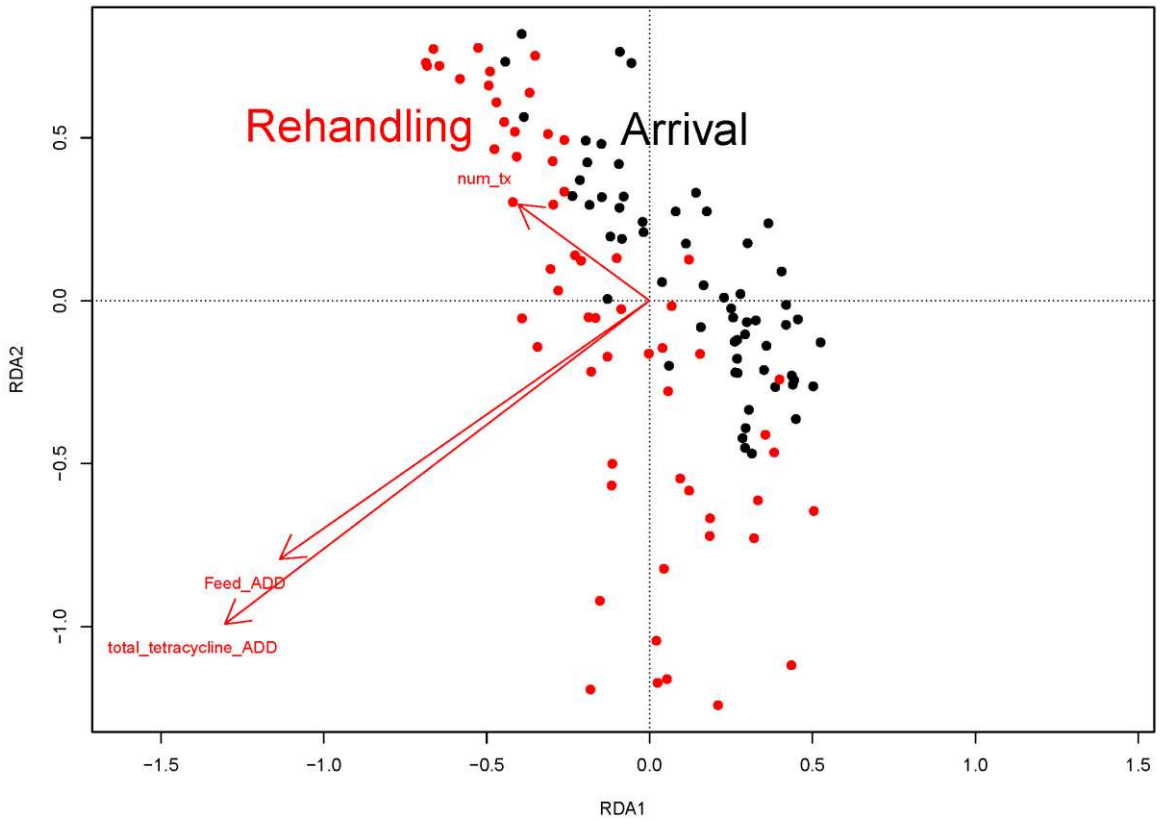


Figure 4.7. Redundancy analysis (RDA) plot illustrating the relationship between microbiome composition and antimicrobial drug exposure metadata. Step wise model selection was used to determine the combination of variables that best explain the variance in the microbiome. Sampling time and in-feed exposure to antimicrobial drugs (in ADDs) were included in the final model and statistically associated with explaining 1.2% and 0.2% of the microbiome variance, respectively.

CHAPTER 5: METAGENOMIC INVESTIGATION OF THE EFFECTS OF ANTIMICROBIAL DRUG USE PRACTICES ON THE MICROBIOME AND RESISTOME OF BEEF FEEDLOT CATTLE

INTRODUCTION

Beef production systems utilize antimicrobial drugs (AMDs) to greatly benefit animal health and welfare, which in turn improves farm sustainability, but the use of these AMDs are under increasing scrutiny regarding the risk of antimicrobial resistance (AMR) that affects treatment efficacy in humans (15, 169). The presence of AMR pathogens, has been shown to affect the duration of morbidity, risks for mortality, need for intensive alternative treatments, increased duration of hospitalization, and costs of treatment (174). Genes conferring AMR are a naturally occurring part of microbial ecology, which were present in bacteria for several millennia prior to the discovery and use of AMDs by humans (2, 4). Concern is mounting, however, that the increasing use of AMDs in health care and agriculture is associated with an increase in the prevalence of AMR among pathogens (175–177), but there is poor understanding of how changing AMD use practices can reduce the risk of AMR. Notably, beef production uses AMDs from drug classes considered important for human medicine (12) and the concern is that this will directly select for increased resistance to these drugs which can then be disseminated into the environment (10, 178, 179). To balance the use of AMDs and ameliorate the risks of AMR promotion and transmission from livestock production systems, it is critical that we understand how different AMD use practices in beef feedlot production affect AMR dynamics.

Until the last decade, studies regarding AMR were primarily culture-dependent and consisted of using aerobic culture to isolate a few bacterial species and test their resistance

patterns with phenotypic testing (i.e. broth microdilution, Kirby-Bauer disk diffusion). These studies have unavoidably targeted culturable bacterial species such as a few bacterial pathogens (e.g., *Campylobacter* spp., *Salmonella enterica*, etc.) and indicator species (e.g., *Escherichia coli*). While estimating AMR prevalence among a specific limited set of bacterial species can be used for evaluating trends in resistance over time, we are lacking the ecological perspective of how AMR is transmitted among all constituents of the microbial community (microbiome) in the environment. High-throughput genetic sequencing now allows for the culture-independent characterization of both the microbiome and resistome. This approach will help to expand our understanding of how bacterial communities respond to AMD exposure in beef feedlot production systems.

Our study leverages archived fecal samples collected in a previously described longitudinal study conducted in four Canadian feedlots with detailed AMD use records (34, 171). In that study, newly formed pens of feedlot cattle were randomly selected for inclusion into the study and composite pen-floor fecal samples were collected at two time points during the feeding periods. Pens in the study were managed under conditions typical at large commercial beef feedlots in North America, and the AMD use records regarding individual treatment of animals and in-feed inclusion of AMD were aggregated to estimate the selection pressure imposed on the resistome and microbiome. The gastrointestinal commensal species, non-type-specific *Escherichia coli,* was cultured from all samples and tested for AMR using broth microdilution and Kirby-Bauer disk diffusion. The goal of this study is to use next-generation sequencing on those samples to investigate the effects of AMD use on AMR dynamics in the microbiome and resistome.

METHODS

Study design

In this study we employ metagenomic sequencing to characterize the effect of AMD use on the microbiome and resistome in feces collected during a previously published 3-year longitudinal study of Canadian beef feedlot operations. Out of all pens of cattle arriving to the feedlot, 30% were randomly selected for inclusion into the study. All exposure to antimicrobial drugs, including parenteral treatments and in-feed exposures, was recorded and standardized across different drug classes using animal defined daily dose (ADD). For each pen in the study, the ADD exposure for all animals within that pen were aggregated to pen-level ADD exposure. Fecal samples were selected from archived samples processed for 16S rRNA sequencing and shotgun metagenomics with an AMR gene bait-capture system to describe the microbiome and resistome, respectively.

Sample Collection

Samples were collected from four participating feedlots in western Canada and composite pen-floor fecal samples were collected from each pen soon after arrival (Arrival samples) and at a second time point (Re-handling) > 60 days on feed (DOF), as previously described (171). Samples were collected longitudinally from 8 October 2009 to 6 April 2010 and stored in Cary-Blair transport medium as previously described (34). Samples were archived in Whirl-Pak bags or 30mL medium transport tubes at -80°C until shipment on dry ice to Colorado State University in May of 2017. They were stored at -80°C again until samples were individually thawed for DNA isolation and extraction. Each Arrival sample, Re-handling, and Shipment sample was

labeled and archived together, so that these corresponding samples from each pen could be correlated after extraction.

Sample selection for DNA Isolation and metagenomic sequencing

Pen samples were selected using stratified random-sampling to include pens with Re-handling samples collected before and after 100 DOF. From each DOF time group, 19 pen samples were randomly selected. Further, a group of 6 pens with the highest ADD at Re-handling were purposively selected for inclusion. A total of 44 pens were selected for enrollment into this study and fecal samples collected from two time points were included for DNA isolation and metagenomic sequencing. Additionally, 10 samples were included from pens sampled a third time just prior to shipment to the abattoir. Fecal samples were processed for DNA extraction and metagenomic sequencing using laboratory methods as described in Chapter 4: "Antimicrobial Drug Use in Beef Feedlots; Effects on the Microbiome and Resistome Dynamics in Individual Cattle". Briefly, with the exception of increasing centrifugation time from 3 to 5 minutes, we utilize the Qiagen PowerMax Soil Kit (Qiagen Laboratories) according to manufacturer's instructions. DNA precipitation by ethanol increases DNA concentration and can be quantified using the Qubit 2.0 Fluorometer and its corresponding dsDNA High Sensitivity Buffer and Reagent (Thermo Fischer Scientific). Quality and sample purity are further verified using NanoDrop 1000 Spectrophotometer (Thermo Fischer Scientific). Pooled samples (30μL) with an average fragment length of 292 bp were shipped on ice to the Novogene Corporation for 16S rRNA amplicon sequencing. Library preparation was performed in-house for resistome characterization and prepped libraries were delivered to UC-Denver Genomics and Microarray Core Facility (Denver, CO) for shotgun sequencing.

Library Preparation and Metagenomic sequencing

For microbiome sequencing, between 200-500 ng of DNA from each sample (N=98) was sent to the Novogene Corporation for 16S rRNA gene amplification and sequencing on the Illumina HiSeq 2500 Sequencing System. The primer set 515F/806R [5'-GTGCCAGCMGCCGCGGTAA-3']/[5'-GGACTACHVGGGTWTCTAAT-3'] was used to amplify the V4 region of the 16S rRNA gene and sequence paired end 250 bp reads (PE 250) at a targeted read depth of up to 100,000 PE reads per sample. Following total DNA extraction, sample aliquots were prepared for resistome sequencing using the SureSelectXT HS Reagent Kit (Agilent Technologies) combined with a custom bait design targeting AMR genes, 'MEGaRICH' (84). Sequencing libraries were transported to UC-Denver Genomics and Microarray Core Facility (Denver, CO) and sequenced using the NovaSeq 6000 Sequencing System (Illumina), targeting a read depth of 100 million PE reads per sample and paired end 150bp reads (PE 150). Systematic bias in sequencing results was assessed across the following sequencing metadata: sequencing run, batch, and lane. Generalized linear models were fit on the number of raw, trimmed, and non-host reads using the R programming language (126) and the "glm" function.

Microbiome and Resistome characterization

To describe the profile of microbes present in feedlot pen floors, reads from 16S rRNA amplicon sequencing were analyzed using the collection of tools contained in Quantitative Insights Into Microbial Ecology version 2 (172). Briefly, all reads are processed for sequence quality and denoising using DADA2 (180), taxonomic classification using a naive bayes classifier trained on the GreenGenes database (97), and removal of chloroplast and mitochondrial

DNA contaminants. Results were then exported into count tables and summarized using the R statistical software.

Alternatively, to identify the resistome in feedlot pen floors, the targeted AMR metagenomic sequencing samples were processed using the AMRPlusPlus bioinformatic pipeline and the MEGARes resistance database (49). Further details on the pipeline can be found in the documentation website (http://megares.meglab.org/amrplusplus) and have been previously reported (56). Sample read trimming, quality filtering and host DNA removal are all performed to create "non-host reads" that are used for subsequent analysis of the resistome using AMRPlusPlus. The AMRPlusPlus pipeline now accounts for PCR duplication by de-duplicating alignments to MEGARes and improves the classification of AMR genes requiring the presence of specific SNPs by incorporating Resistance Gene Identifier (95). The "strict" paradigm was employed as this incorporates detection models and CARD's curated similarity cut-offs to increase accuracy in identifying functional AMR genes (see full script: https://github.com/EnriqueDoster/MEG-amrplusplus-RGI-SNPconfirmation). A list of important AMR gene determinants that are considered to be critically important to public health when identified in human pathogens were specified *a priori* and sequences were specifically searched for in all samples (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat, cfr).

Count matrix processing

The R statistical programming software and publicly available packages, "metagenomeSeq" and "vegan" (130, 131) were used to analyze count tables representing the resistome and the microbiome. Raw counts of alignments were normalized to account for

differences in sequencing depth using Cumulative sum scaling (CSS) (131). Sparsely represented features identified in fewer than 5% of samples were removed from further analysis based on published recommendations (131). Resistome counts were summed to the drug class and resistance mechanisms level for each sample to avoid bias at the "gene" level associated with irregular naming criteria for new resistance genes (132). Similarly, microbiome counts were aggregated to these 6 Linnaean taxonomic levels: phylum, class, order, family, genus, and species. In all, 6 count matrices for the microbiome, but to reduce the repetitive reporting of results at all levels and because results at lower taxonomic levels are not considered reliable (133), statistical results for microbiome are presented at the phylum, class and order levels. Diversity was calculated for each sample using "vegan" and summarized with richness and Shannon's diversity index.

Statistical analysis

The primary exposure variables of interest were time in feedlot, as evaluated by comparisons of microbiomes/resistomes collected at the two sampling times (Arrival vs Re-handling) and magnitude of ADD exposures. ADD exposures were summarized for each pen by calculating the sum of ADDs from parenteral and in-feed treatments grouped into 3 exposure categories; Low ADD exposure (< 1000 total ADD), medium ADD exposure (1000 – 5000 total ADD), and high ADD exposure (> 5000 total ADD). Similarly, the DOF for each animal at the time of sample collection was categorized into 5 ranges (arrival - 3 DOF, 4-70 DOF, 71-120 DOF, 121-180 DOF, and > 180 DOF). Diversity indices were statistically compared using the Wilcoxon signed-rank test ("wilcox.test" function in R) for samples from the same animal and "glm" to test differences between other sample groups. CSS- normalized counts were Hellinger-

transformed (135) for ordination using the metaMDS function from “vegan”, which employs non-metric multidimensional scaling on Euclidian distances. Analysis of similarities (ANOSIM) (136) was used to test differences in the microbiome and resistome between categorical metadata sample groups (e.g., Arrival vs Re-handling samples, ADD exposure category, and DOF sampling category. MetagenomeSeq’s “fitZig” function was used to fit a zero-inflated Gaussian model and compare log2-fold differences (131) in microbiome and resistome features between sampling time. Limma’s “makeContrast” function (137) were then used for pairwise comparisons, P-values were adjusted for multiple tests using the Benjamini-Hochberg procedure (138), and alpha=0.05 was selected as the statistical significance cut-off value. To account for spurious statistically significant differences in low abundance features, only features with an average expression > 1 were considered.

Additionally, raw counts were hellinger-transformed (135) and redundancy analysis was performed on the microbiome and resistome composition to further evaluate the potential significance of different AMD exposures using the “rda” function in R. The significance of the correlation between independent variables and the variance in the microbiome and resistome composition were then tested using the “anova” function in R. With the goal of characterizing the effect ADD exposure and time in the feedlot (DOF) have on the microbiome and resistome, samples were grouped into 18 metadata categories for analysis (Supplemental Table 5.1). ADD exposures were aggregated by route of administration (in-feed vs parenteral) and by drug class including macrolides-lincosamides-streptogramin (MLS), tetracyclines, phenicol, and sulfamethoxazole-trimethoprim. Categorical and numerical categories were created to summarize the amount and type of antimicrobial drug exposure as well as time in the feedlot and days since the most recent parenteral treatment. All of these variables were included in the

starting model for step-wise variable selection and ANOVA testing to identify the best fitting model. Additionally, a model was evaluated with *a priori* inclusion of variables regarding the duration that cattle had been in the feedlot environment (days on feed or DOF), sum of ADD exposures in feed, and the sum of parenteral ADD exposure.

RESULTS

Sequencing results

In all, samples from 44 pens of cattle were selected for metagenomic sequencing and included two sampling time points for all samples (Arrival vs Re-handling) and a third sampling point for 10 samples for a total of 98 samples. Sequencing of the resistome using target-enriched shotgun sequencing produced 1.61 billion paired-end reads across 98 samples (mean: 16,470,077 reads per sample, range: 6,192,389 - 25,456,702 reads). Of these reads, filtering to improve overall read quality and to remove bovine host DNA excluded a total of 301.7 M reads, and on average, these reads made up 18.7% of each sample (range: 3.3% - 48% [Supplemental Table 5.2]). There was a lower number of raw sequencing reads in samples from the 10 samples collected immediately prior to shipment (mean = 14,429,269 reads per sample) compared to arrival samples (mean = 17,101,05 reads per sample) and samples collected at the second time point (mean = 16,302,917 reads per sample; $P < 0.05$). However, this difference in number of sequencing reads was not present after removal of contaminant host DNA. Comparably, there were no differences in sequencing effort for the microbiome. More than 19.4 million reads were produced from 16S rRNA sequencing across 98 samples (mean: 197,889, range: 94,433 – 219,918). Quality filtering, identification of sequence variants with DADA, and removal of chloroplast and chimeric sequences identified a total 4,941,757 sequence variants with each sample averaging 50,426 unique sequence variants per sample (range: 25,435 – 81,128).

Resistome results

Following alignment of reads to the MEGARes AMR database, removal of duplicate reads, and re-classification of reads aligning to gene accessions requiring SNP confirmation with RGI, a total of 55,155,938 reads aligning to AMR gene accessions ("hits") were identified across 98 samples (mean: 562,816 reads per sample, range: 278,635 – 1,084,706). These hits were classified as corresponding to 1,951 different gene accessions which represented genes that confer resistance to 26 different AMD classes through 91 distinct resistance mechanisms. Out of the 26 drug classes identified, hits to a majority (19/26) were present in low abundance and together comprised only 1.4% of all reads aligning to AMR gene accessions. After removal of sparse features and CSS normalization of counts, hits to the seven most abundant drug classes were represented by tetracyclines (52.4%), aminoglycosides (11.2%), MLS (10.8%), mechanisms conferring resistance to multiple types of drugs (e.g., multidrug efflux pumps -- 9.6%), betalactams (6.4%), sulfonamides (5%), and phenicol (3.2%) with the remaining 18 classes each consisting of <1% of normalized counts. Of the hits aligning to genes that confer tetracycline resistance, 68.4% were aligned to gene accessions encoding tetracycline resistance ribosomal protection proteins and 28.6% encoded for major facilitator superfamily (MFS) efflux pumps. In the second most abundant resistance class, aminoglycosides, the 3 most abundant resistance mechanisms were aminoglycoside O-phosphotransferases, aminoglycoside O-nucleotidyltransferases, and aminoglycoside efflux pumps.

Tetracycline and multi-drug resistance mechanisms made up the majority of resistome composition in arrival samples, but by the second sampling period aminoglycoside resistance had replaced multi-drug efflux pumps as the second most abundant drug class (Figure 5.1**)**. ANOSIM testing suggests that the overall resistome composition shifted significantly over time

at the class (ANOSIM R = 0.12, P = 0.001) and mechanism levels (ANOSIM R = 0.13, P = 0.001) (Figure 5.2). Despite these changes in resistome composition, richness and Shannon's diversity indices did not undergo a significant change over time at any resistance level. To identify which resistance determinants in feedlot pens shifted in relative abundance over time, a ZIG model was fit on 23 drug classes and identified that 15 drug classes were present at an average expression > 1, and the only significant changes occurred to 9 drug classes decreasing from arrival to the second time point samples ($P < 0.05$). Similarly, 21 mechanisms shifted in abundance and only 1 mechanism, Macrolide resistance efflux pumps, increased between sampling time.

To further investigate how much of the variation in the fecal resistomes of feedlot cattle could be explained by independent variables (sampling time, feedlot ID, AMD exposure in ADD units, etc.), redundancy analysis was performed starting with a full model containing all 14 independent variable categories (Supplemental Table 5.1). Step-wise selection to identify the best model included 4 independent variables (Sampling Time, parenteral macrolide ADD, total ADD, and parenteral fluoroquinolone ADD), and explained 1.4% of the constrained variance and 4.4% of the total variability among resistomes in different samples (Figure 5.3.A**)**. Of the 4 independent variables in the model, only sampling time and total ADD were statistically significant (ANOVA $P < 0.05$) representing 0.9% and 0.3% of the constrained variance, respectively. Because of the significant shifts in resistome composition over time, samples collected at the second time point and just prior to shipment were found to be highly similar in the redundancy analysis. For these samples collected after arrival, the redundancy model with the best fit included 3 variables (ANOVA $P < 0.05$): total parenteral fluoroquinolone ADD exposure, in-feed tetracycline ADD exposure, and feedlot ID (Figure 5.3.B). The final model explained

2.6% of the variability in the resistomes of different samples, and suggested that only 0.77% of that variation can be explained by the independent variables measured in this study. Of the constrained variance, total parenteral fluoroquinolone ADD exposure and feedlot ID were significantly associated with explaining only 0.33% and 0.38% of the variance in the resistome, respectively ($P < 0.05$).

Of the a *priori* list of genes selected as being important to public health when expressed in human pathogens, the following gene groups were identified in at least one sample in our study: (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(GES), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat(A/B/C/D/E), cfr). Across 98 samples, 307,485 reads mapped to these genes and were predominantly (>95.6%) comprised of reads aligning to bla(CTX-M), bla(OXA), bla(TEM), cfr, and vat(A/E) gene accessions. Between sampling times, the number of samples (each representing different pen populations) containing reads aligning to these 6 gene accessions generally was similar, whereas the number of samples with low abundance gene groups decreased over time (Figure 5.4).

Microbiome results

A total of 4,328,330 reads were classified taxonomically into 29,683 amplicon sequence variants. Overall, these taxa represent 25 phyla, 73 classes, and 113 unique taxonomic orders. Five phyla, Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, and Tenericutes predominated in the microbiome and accounted for > 95.6% of all normalized counts (46.5%, 21.1%, 20.1, 5% and 4.9% respectively). The top three phyla consisted of a single taxonomic order making up a majority of counts. Across all samples, the order Clostridiales made up 73.2%

of Firmicutes alignments, the Bacteroidales order made up 97.8% of Bacteroidetes, and Pseudomonadales were the dominant order of proteobacteria (97.3%).

Overall, the microbiome composition in feedlot pen floor fecal samples was dominated by Firmicutes, Bacteroidetes, and Proteobacteria at all sampling time points with the majority of phyla (20/25) accounting for < 0.01% of all classified reads (Figure 5.5). There were no significant changes in Shannon's diversity index at any level. At arrival the microbiome richness averaged 11.5 at the phylum level, 20.2 at the class, and 27.3 at the order level, but richness decreased significantly from Arrival to Re-handling samples at the class (18.6) and order (25.3) levels ($P < 0.05$). Further, significant shifts in the microbiome were observed at the phylum (ANOSIM $R = 0.08$, $P = 0.01$), class (ANOSIM $R= 0.12$, $P < 0.01$), and order (ANOSIM $R= 0.14$, $P < 0.01$) taxonomic levels between sampling times (Figure 5.6). Eleven phyla were identified with an average expression >1 and 6 phyla significantly changed in relative abundance from Arrival to Re-handling. Cyanobacteria, Verrucomicrobia, and actinobacteria decreased significantly in abundance over time accompanied with increases in Tenericutes, Spirochaetes, and Euryarchaeota ($P < 0.05$).

Redundancy analysis and step-wise selection of metadata variables that best explain the variance in the microbiome resulted in a statistically significant model (ANOVA $P < 0.05$) explaining 8% of the variance between all samples (Figure 5.7.A). Three variables, DOF, total parenteral ADD exposure, and parenteral sulfonamide ADD were included in the model together explained 1.3% of the constrained variance. While all three variables were significantly associated with the microbiome composition ($p < 0.05$), results suggest DOF explains a higher variance (0.87%) than total parenteral ADD exposure (0.25%) or parenteral sulfonamide ADD (0.2%). Performing redundancy analysis on all other samples not collected at arrival resulted in a

statistically significant model (ANOVA $P < 0.05$) explaining 5.3% of the variance between all samples. Feedlot ID, total in-feed ADD, DOF, and parenteral fluoroquinolones ADD were included in the model, but only feedlot ID and total in-feed ADD were statistically significant representing 0.63% and 0.36% of the microbiome variance ($P < 0.05$) (Figure 5.7.B).

DISCUSSION

Our results suggest that exposure to antimicrobial drugs has a relatively small influence on the microbiome and resistome of feedlot pens compared to shifts in composition associated with DOF after arrival to the feedlot. Overall, our redundancy analysis could only explain < 9% of the variance in the microbiome and resistome which suggests that other un-measured factors in the environment likely impose a greater overall impact.

The resistome and microbiome in all samples were largely dominated by a few abundant features, namely alignments to AMR gene accessions that confer resistance to tetracycline, aminoglycoside, and MLS drug classes in the resistome and the microbiome consisting of Firmicutes, Bacteroidetes, and Proteobacteria. Of all tested metadata variables, the largest measurable association was sampling time (Arrival vs Re-handling) and could account for 0.9% of the variation in the resistome compared to 0.29% explained by total AMD exposure (ADDs), but these changes in the resistome were not associated with increasing DOF when only considering the second sampling point. In samples collected from pens > 60 DOF, a small and statistically significant percentage of the variance in the resistome could be explained by the originating feedlot and increasing use of parenteral treatment with fluoroquinolone AMDs. The microbiome likewise was significantly influenced primarily by sampling time and by the amount of parenteral AMD exposure to animals in the pen. These major short-term shifts in microbial populations occur as cattle are introduced to the feedlot environment (56, 173) and our results

suggest that by later in the feeding period, observed differences in the resistome particularly in low abundance features can be due to different AMD use practices in distinct feedlots.

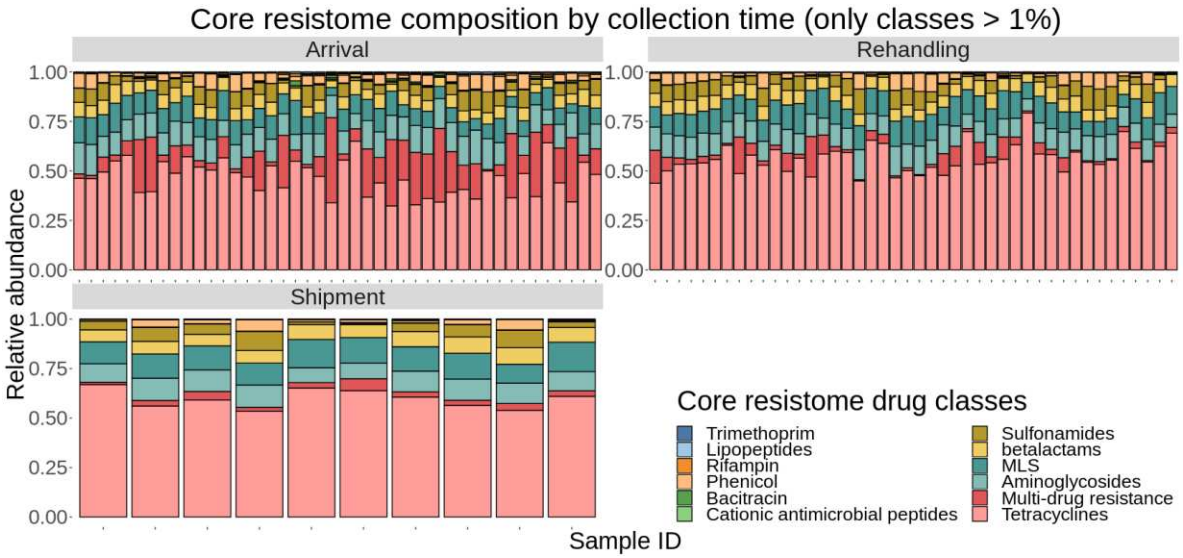
Alignments to the MEGARes database were interrogated for matches to an a-priori list of genes with medical importance. We identified 16 a-priori genes across 98 samples, but only the following six gene groups were found in > half of the samples at the second sampling time and accounted for >98% of alignments to the a priori gene accessions; cfr, bla(OXA), bla(CTX-M), bla(TEM), and vat(A/E). Though samples were collected from 44 different pens in 4 commercial feedlots, similar patterns emerged in the prevalence of these genes over time in the feedlot. The bla(CTX-M), bla(OXA), and bla(TEM) group of gene accessions were in the highest relative abundance of important AMR genes and could be identified in 97.7% (43/44) of all samples at arrival. Their high prevalence was maintained by the second sampling time point with only bla(OXA) and bla(TEM) found in one less pen sample. Though the low abundance of normalized counts precluded the ability to statistically compare logFold changes in abundance over time, the sum of alignments decreased for all 6 of the genes identified. Importantly, the function of all these medically important AMR gene accessions cannot be inferred from the methods employed in this study and the bacterial species carrying these AMR genes cannot be determined. Similarly, classification of AMR gene accessions in our samples does not ensure the functional activity of these genes and instead serves only to identify the potential resistome function. Further, sequencing depth could have been inadequate to identify more subtle dynamics occurring in low abundance or “rare” features (181). Therefore, differential expression of AMR genes cannot be accurately assessed, and future studies will require incorporating multiple techniques to better characterize the risk of AMR emergence.

The overall strategy employed in this study to use archived samples and re-process those samples for metagenomic sequencing is an example of what is possible with high throughput sequencing. Previously collected samples can be re-used to add new community-level insights to results originally limited by traditional methods. Beyond being used to study AMR dynamics in concurrent studies, metagenomic sequencing can be employed on high quality samples stored from past experiments to add further community-level insights with a relatively minor added cost. As laboratory methods for DNA extraction and sequencing are known to bias sequencing results, future studies must consider the external validity of microbiome and resistome results garnered from samples not originally processed for DNA extraction and sequencing (77). Originally, the samples utilized in this study were mixed with Cary Blair media and stored in a refrigerator prior to freezing. Therefore, our internal validity is sound as all samples were processed with the same protocol, but comparisons to external studies should be done with caution.

To the author's knowledge, no other metagenomic projects have been performed on composite-pen floor samples of commercial feedlots to this scale. Use of high-throughput sequencing for AMR surveillance in feedlot environments is set to revolutionize our understanding of AMR dynamics, but this is reliant on the availability of ample details regarding parenteral and in-feed use of AMDs along with environmental factors about the sample collected such as temperature, moisture level, pH, and geographic location (102, 182, 183). Through a better understanding of which factors influence the microbiome and the resistome in the feedlot environment, AMD use practices can be better tailored to reduce the risk of promoting the emergence of AMR in bacteria and the environment. To eventually learn how to manage AMR in

a commercial feedlot, as sequencing costs continue to decrease sampling pen-floors may provide a cost-effective manner for surveillance of community wide trends.

A)



B)

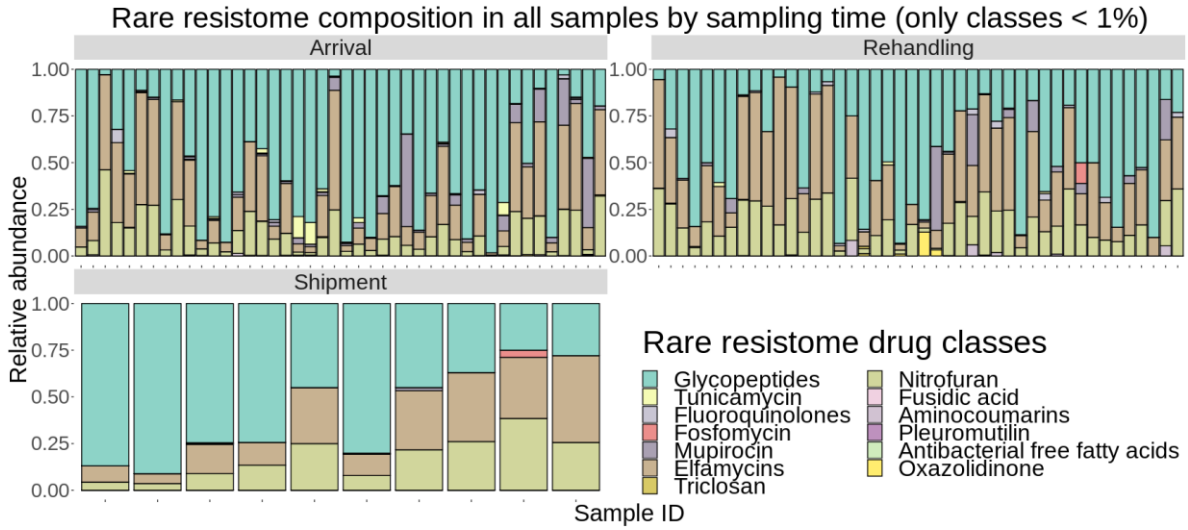


Figure 5.1. Resistome composition at the drug class level for all samples by sampling time point; arrival at feedlot, a second time point during re-handling of pen cattle, or just prior to shipment to the abattoir. A) Shows the 12 drug classes present in a relative abundance > 1% (Core) of all resistome alignments and B) is the subset of 13 drug classes in low relative abundance (Rare).

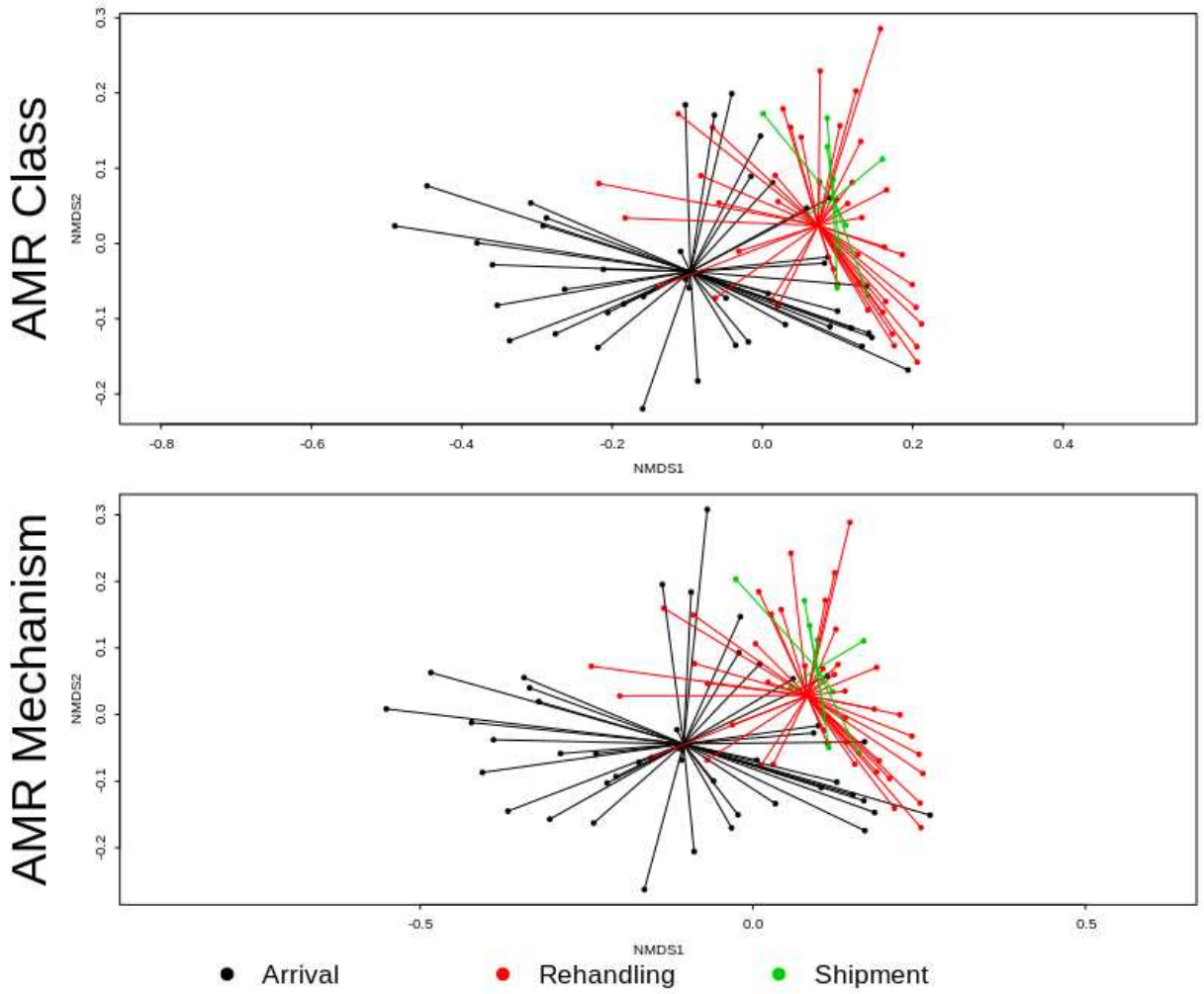
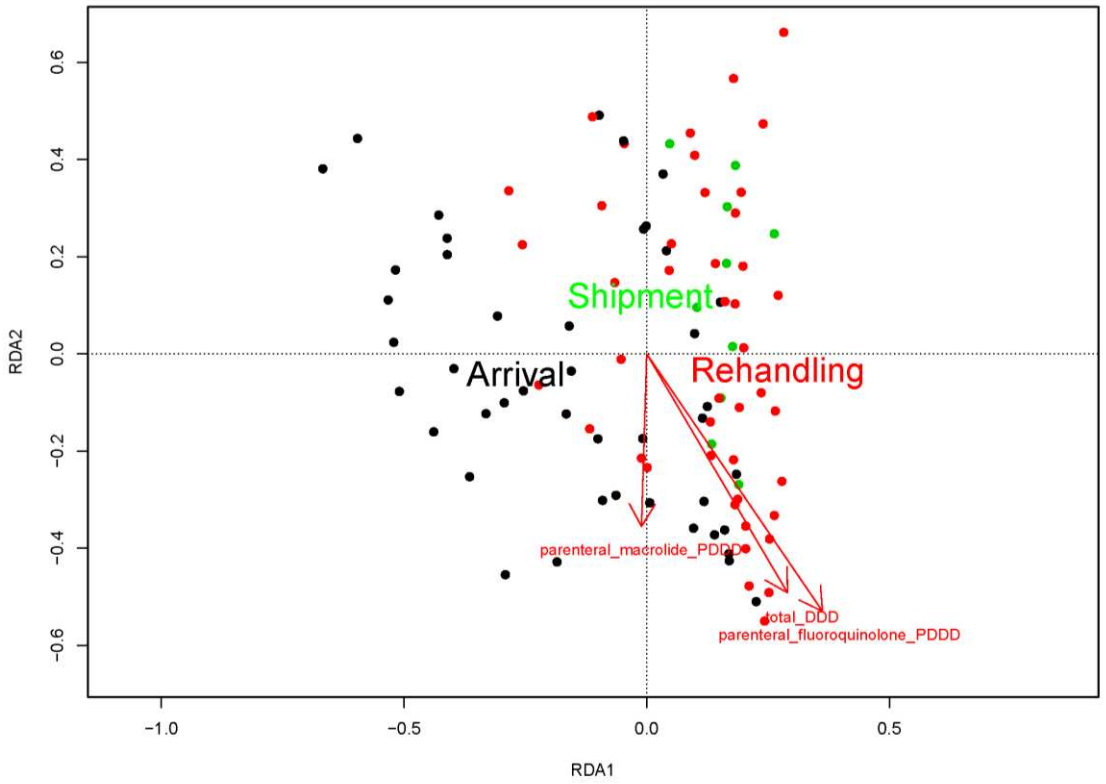


Figure 5.2. Ordination comparing resistome composition at the AMR drug class and resistance mechanism, using non-metric multidimensional scaling (NMDS), for samples collected at Arrival, Re-handling, and just prior to Shipment. Separation of resistomes between sampling time was statistically significant at the class (ANOSIM R = 0.32, $P < 0.05$) and mechanism levels (ANOSIM R = 0.34, $P < 0.05$).

A)


Shipment
Arrival
Rehandling
parenteral_macrolide_PDDD
total_DDD
parenteral_fluoroquinolone_PDDD
RDA1
RDA2

B)

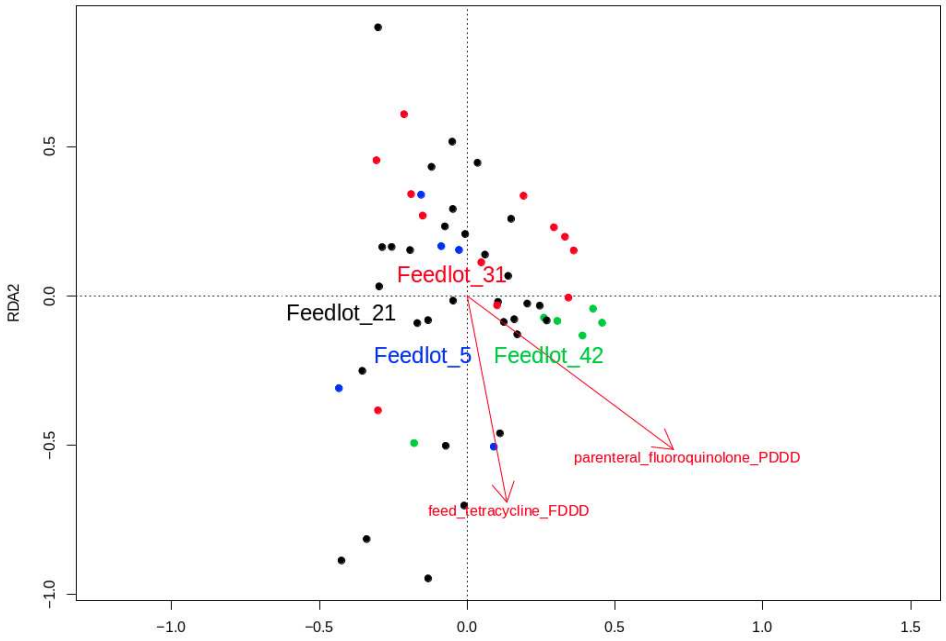
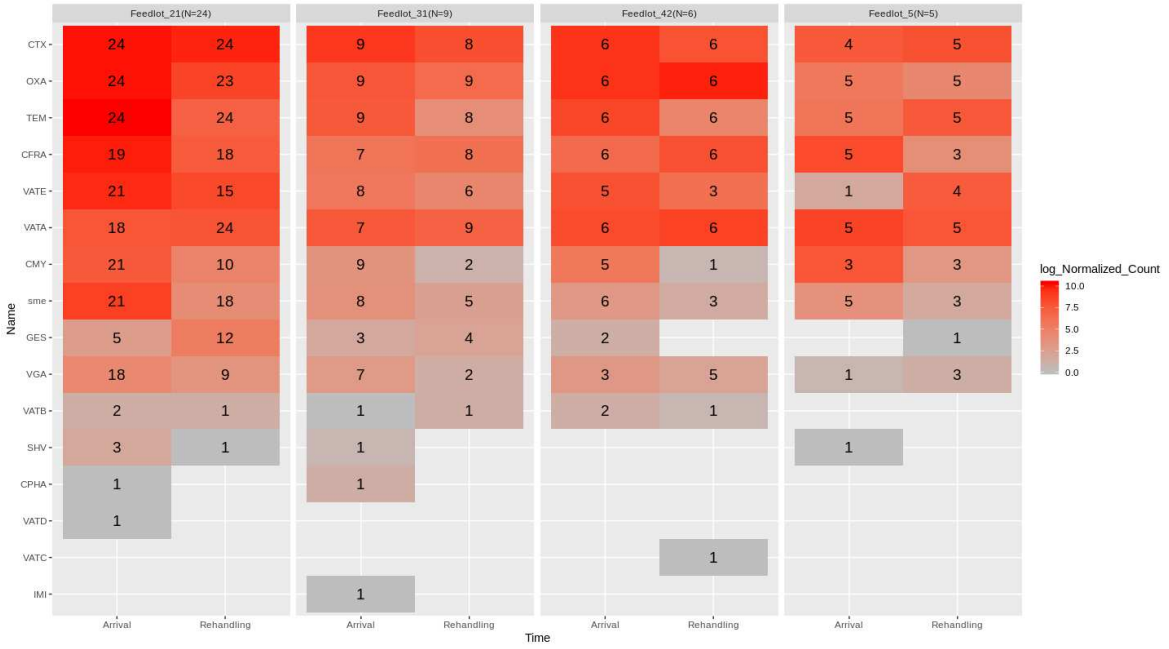

Feedlot_31
Feedlot_21
Feedlot_5
Feedlot_42
parenteral_fluoroquinolone_PDDD
feed_tetracycline_FDDD
RDA1
RDA2

Figure 5.3. Redundancy analysis (RDA) plot illustrating the relationship between resistome composition and antimicrobial drug exposure metadata. Step wise model selection was used to determine the combination of variables that best explain the variance in the microbiome. A) Considering all samples, Time, parenteral macrolide ADD, total ADD, and parenteral fluoroquinolone ADD were included in the model and explained 1.4% of the constrained variance. B) Excluding arrival samples and performing the same methodology resulted in a model with total parenteral fluoroquinolone ADD exposure, in-feed tetracycline ADD exposure, and feedlot ID included in the model. Parenteral fluoroquinolone ADD exposure and feedlot ID were significantly associated with explaining 0.33% and 0.38% of the variance in the resistome, respectively ($P < 0.05$).

A)



B)

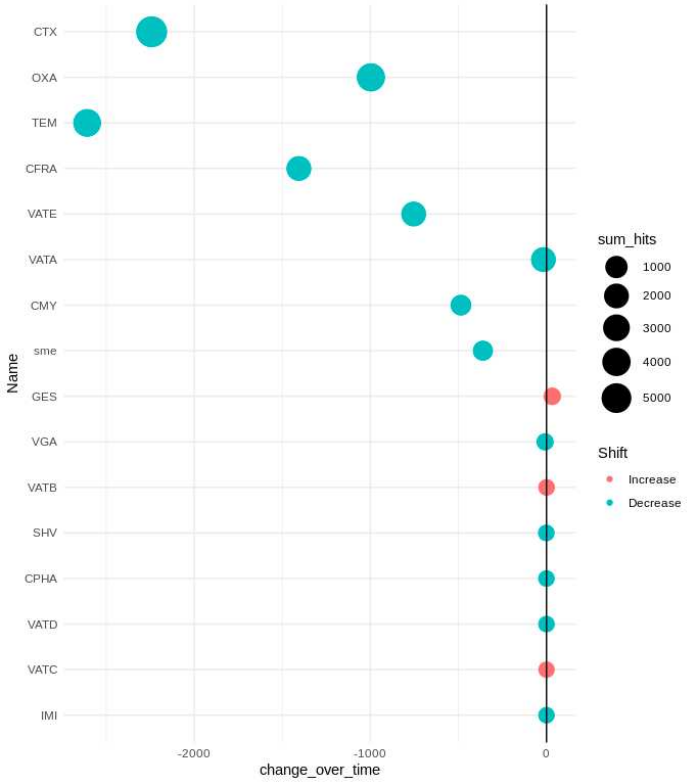
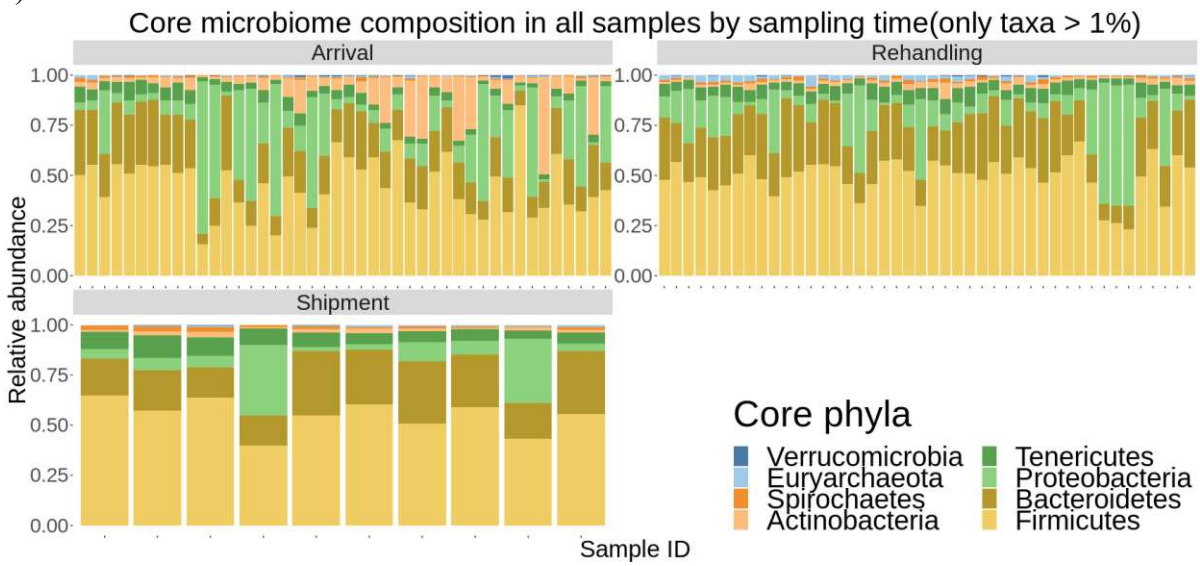


Figure 5.4. A) Heatmap representing the prevalence of medically important AMR genes across 4 different feedlots. The increasing shade of red represents the relative abundance of counts for each gene group on the rows. The number on each cell signifies how many samples contained alignment to that gene group. B) Plot of the change in alignments from the arrival samples to exit for each medically important AMR gene group across all samples.

A)



B)

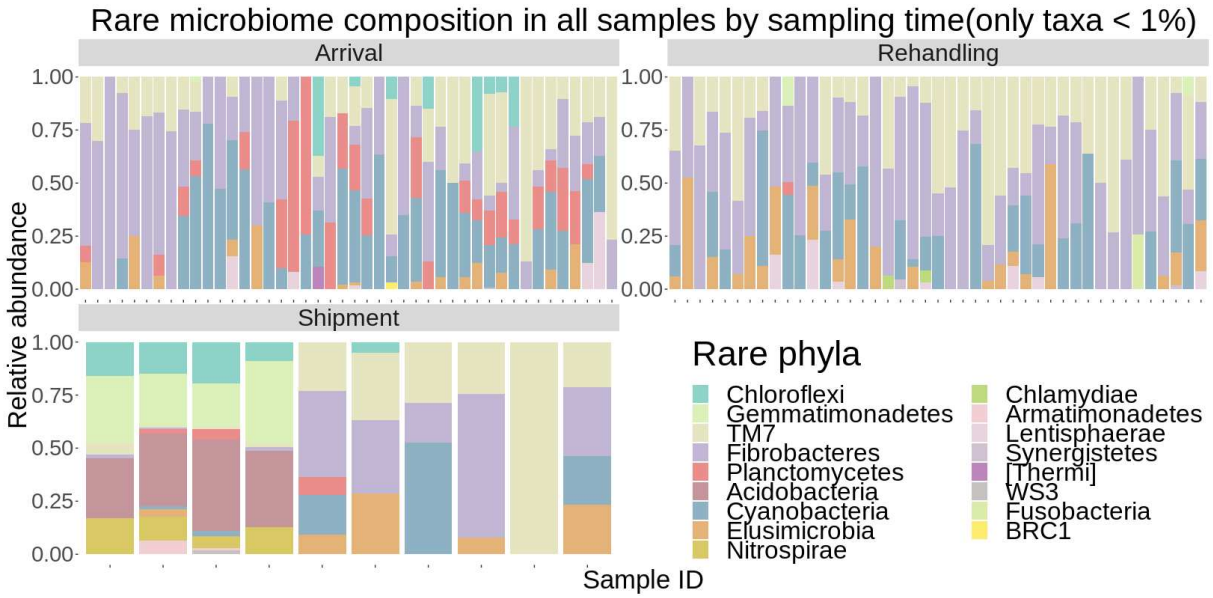


Figure 5.5. Microbiome composition at the phylum taxonomic level for all samples by sampling time point; arrival at feedlot, a second time point closer to exit of the feeding period, or just prior to shipment to the abattoir. A) Shows the 8 phyla present in a relative abundance > 1% of all microbiome alignments (Core) and B) is the subset of 17 phyla in low relative abundance (Rare).

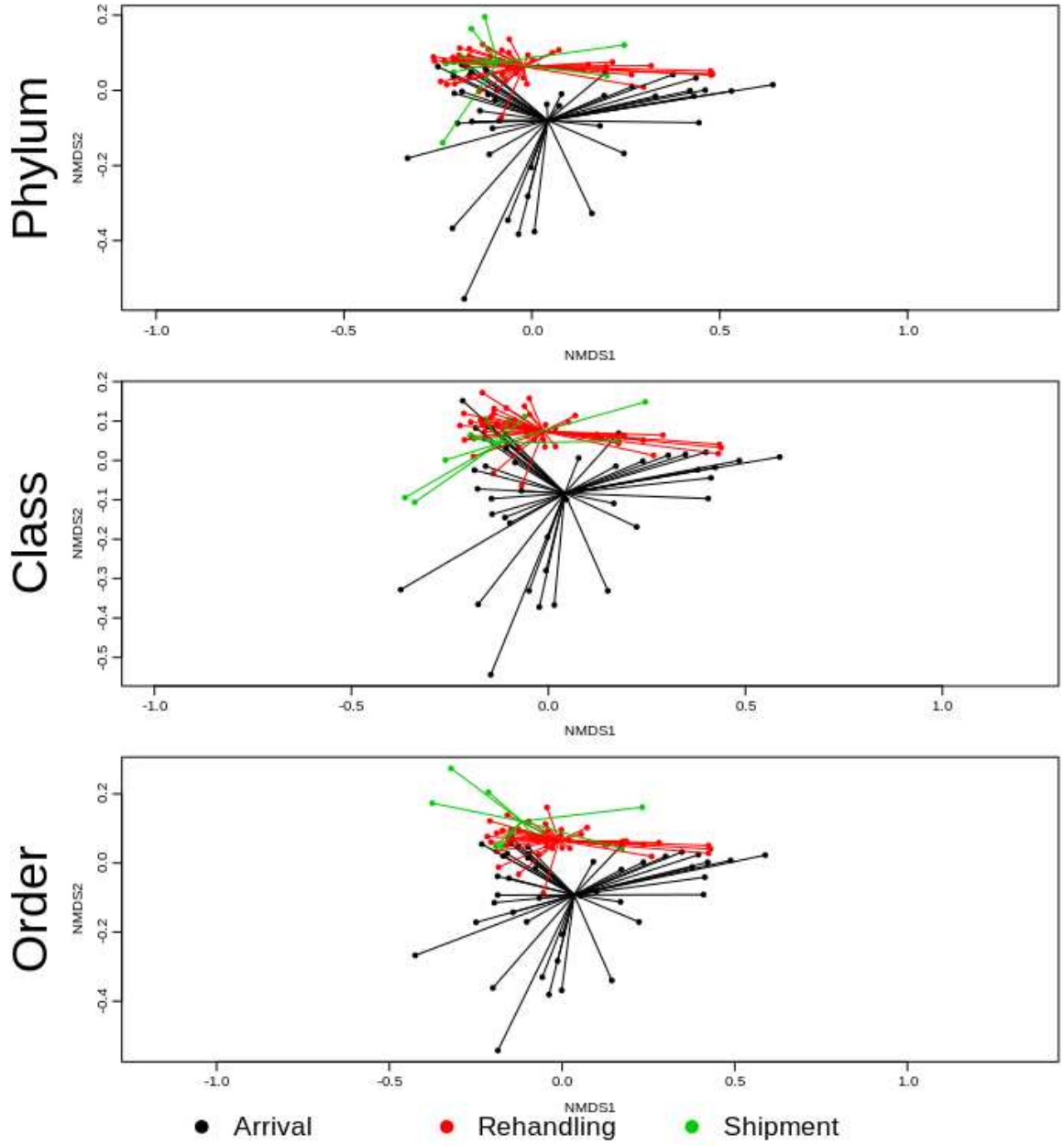
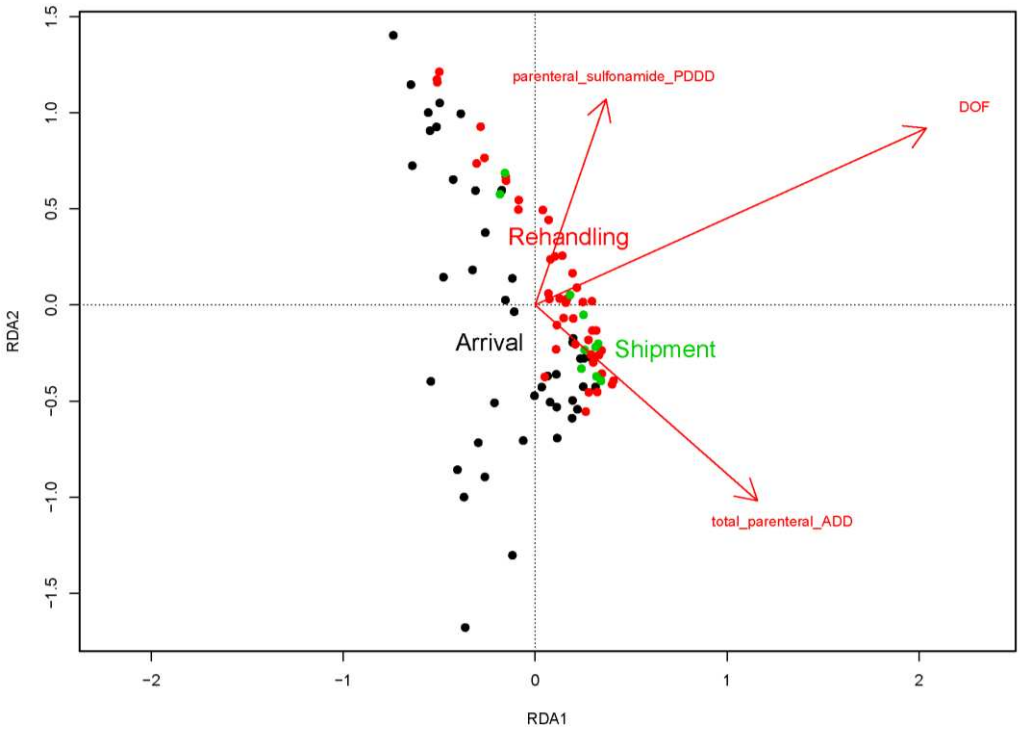


Figure 5.6. Ordination comparing microbiome composition at the phylum, class, and order levels , using non-metric multidimensional scaling (NMDS), for samples collected at Arrival and Re-handling, and just prior to Shipment. Separation of resistomes between sampling time was statistically significant at the phylum (ANOSIM R = 0.08, P = 0.01), class (ANOSIM R= 0.12, P < 0.01), and order (ANOSIM R= 0.14, P < 0.01) taxonomic levels.

A)



B)

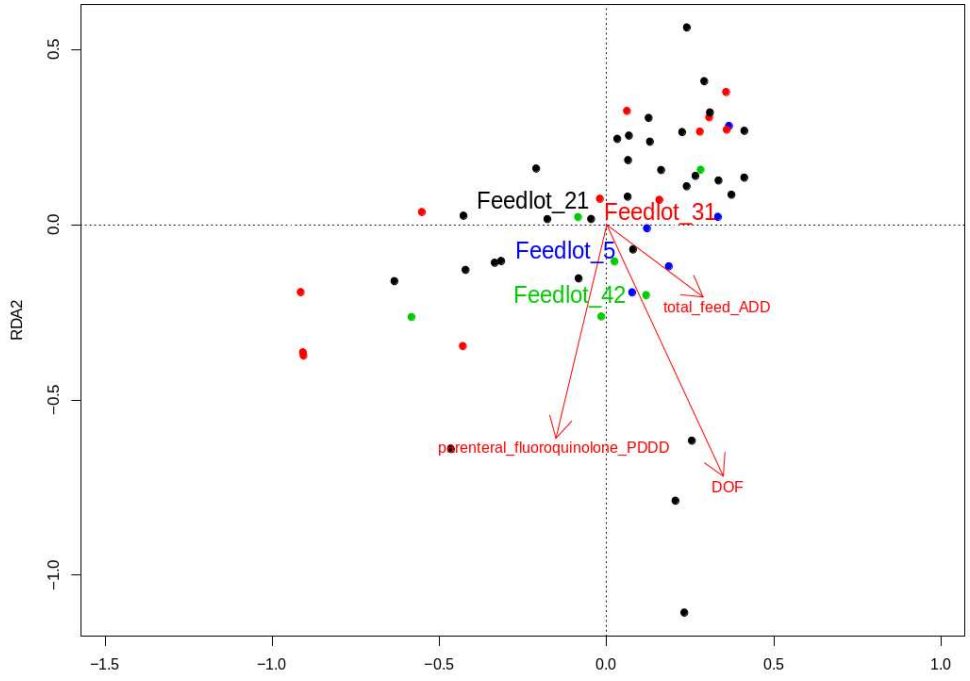


Figure 5.7. Redundancy analysis (RDA) plot illustrating the relationship between microbiome composition and antimicrobial drug exposure metadata. Step wise model selection was used to determine the combination of variables that best explain the variance in the microbiome. A) Considering all samples, DOF, total parenteral ADD exposure, and parenteral sulfonamide ADD were included in the model together explained 1.3% of the constrained variance. B) Excluding

arrival samples and performing the same methodology resulted in a model with feedlot ID, total in-feed ADD, DOF, and parenteral fluoroquinolones ADD included in the model. Feedlot ID and total in-feed ADD were statistically significant representing 0.63% and 0.36% of the microbiome variance, respectively ($p < 0.05$).

CHAPTER 6: METAGENOMIC CHARACTERIZATION OF THE MICROBIOME AND RESISTOME IN RETAIL GROUND BEEF

INTRODUCTION

Ground beef and other foods can be a reservoir for a variety of bacteria, including spoilage organisms and pathogenic foodborne bacteria. Multiple-hurdle intervention systems used at harvest and post-harvest are used to bolster meat safety and have shown to greatly reduce the bacterial load in retail-ready products (184). However, even with ongoing improvements in food safety, the CDC estimates that in 2016, foodborne outbreaks caused an estimated 14,259 illnesses, 875 hospitalizations, and 17 deaths in US. Though beef is not the most common source of foodborne outbreaks, retail ground beef products have been linked to multi-state outbreaks of foodborne illness and it is critical that we understand the bacterial community on retail beef products (65). These bacteria can exhibit antimicrobial resistance (AMR) which is a public health concern as resistance in pathogens can lead to treatment failure in humans. Additionally, concerns regarding of exposure to resistant bacteria or AMDs in food has led to an increase in the practice of marketing animal products as being derived from animals that were "raised without antibiotics" (RWA) to differentiate from products originating from conventional rearing systems (CONV) in which some or all animals were treated with AMDs for control or treatment of disease. Consumers are led to believe that not using antibiotics is associated with decreased risk of exposure to AMR pathogens, but recent studies suggest minimal to no difference in AMR prevalence between these two types of production types (185). Simply removing all antimicrobial use from beef production is not straightforward because of welfare concerns regarding treatment or control of disease. Recent research that there was no significant difference

in the pattern of phenotypic resistance between ground beef product labeled as CONV or RWA (185), little is known about how the production practices could affect the resistome, specially using culture-independent methods for characterizing resistance patterns.

Typically, studies of foodborne bacteria focus on individual pathogens, but high-throughput metagenomic sequencing can now provide a holistic perspective on the community of microbial species (microbiome) and the profile of AMR genes they carry (resistome). Through the use metagenomics, many environments previously considered to be sterile or of low bacterial biomass are now being re-discovered as having a complex microbiome. However, achieving insight into the microbial world comes with challenges in laboratory processing, particularly for food matrices (186). Our research group has developed an approach to target enrich AMR genes present in DNA samples using biotinylated baits. We have demonstrated biotinylated baits ability to increase "on-target" sequencing in cattle feces (MEGARich)(84) and we utilized the same approach to sequence the microbiome and resistome in ground beef samples. The goal of this study was to characterize the microbiome and resistome of retail ground beef products labeled as coming from conventional and raised without antibiotics (RWA) production systems.

METHODS

Study design

Sixteen different ground beef products were purposefully selected and purchased at different retail grocery stores in Fort Collins, Colorado. Samples were processed to allow metagenomic characterization of the microbiome and resistome of conventional (CONV) products and products carrying a "raised without antimicrobials" (RWA) label claim. Total DNA was extracted from each sample and subjected to 16S rRNA amplicon sequencing and target-enriched shotgun sequencing to characterize the microbiome and resistome, respectively.

Microbiome studies of meat samples is complicated due to the high density of DNA percentage and low bacterial biomass. Therefore, this study adds biological replicates for each retail ground beef product to describe intra-sample variability and tests whether a 1 to 2 dilution of baits improves resistome sequencing performance.

Sample Collection

Packages of fresh (not currently frozen) ground beef (≥ 1 lb) were purchased from 6 different retail grocery stores in Fort Collins, Colorado during one day and stored at 4°C until further processed. A variety of product brands and packaging types were purposefully selected in order to represent a broad diversity in sample types and sources. Specifically, different packaging types that were selected included chub (n=4), vacuum-sealed (n=8), and tray wrap (n=4)(187). Additionally, a mix of products were purchased with regard to production claims regarding exposure of source animals to AMDs including products with label claims for certified organic production or specify the lack of antibiotic use during production (n=8), and other products that did not have any label claims specifying antimicrobial drug exposures or organic production claims which were considered to originate from "conventional" production systems (n=8). This characteristic, as well as other metadata such as store name, product brand name, fat percentage, etc. were recorded (Supplemental Table 6.1).

Sample Processing and DNA Isolation

To replicate handling of retail ground beef products by typical consumers, samples were held at 4°C for 48 hours before being opened. Packages were then opened using aseptically by first wiping with 70% ethanol, followed by RNase Away, then and then cut using a sterile

disposable scalpel. Ground beef (30g) was removed from each package and placed in a new Filtra-Bag (VWR). Samples were homogenized using 100 ml of PBS and hand-massaged in a Filtra-Bag. Supernatant (15 mL) was then transferred to a sterile conical tube and centrifuged at 10,000 x g for 10 minutes. The supernatant was then discarded, and pellets were stored at -80°C until processed for DNA extraction. An aliquot of 950 µL of this pellet was used for DNA isolation with the DNeasy PowerFecal Microbial Kit (Qiagen Laboratories) and extracted according to manufacturer's instructions. Isolated DNA samples were eluted in 50 µL of buffer and passed through filter twice to optimize yield. DNA concentrations were measured with the Qubit dsDNA HS Assay Kit using the Qubit 2.0 Fluorometer according to manufacturer's instructions (Thermo Fisher Scientific). If sample concentrations were <1ng/µL, multiple extractions were pooled together to obtain this concentration.

Library Preparation and Metagenomic sequencing

Using between 200-500 ng of DNA from each sample was shipped to Novogene Corporation for 16S rRNA gene amplicon sequencing to characterize the microbiome. The V4 region of the 16S subunit was amplified with the primer set 515F/806R [5'-GTGCCAGCMGCCGCGGTAA-3']/[5'-GGACTACHVGGGTWTCTAAT-3'], with a fragment length of 292 bp. Amplicon sequencing was performed on the Illumina HiSeq 2500 Sequencing System to produce paired end 250 bp reads (PE 250) at a targeted read depth of up to 100,000 PE reads per sample. Additionally, the SureSelectXT HS Reagent Kit for Illumina Paired-End Multiplexed Sequencing Library (Agilent Technologies) was used to prepare samples for target enriched resistome sequencing. A customized bait design targeting AMR genes, 'MEGaRICH' (84), was used to improve "on-target" sequencing and reduce the challenge of sequencing

microbial from a sample predominantly containing host DNA. Consequently, on the recommendation from Agilent Technologies we included a biological replicate of each sample to be processed with ½ diluted baits (N=32) and assess if this improved sequencing performance. Samples were transported to UC-Denver Genomics and Microarray Core Facility (Denver, CO) and sequenced using the NovaSeq 6000 Sequencing System (Illumina), targeting a read depth of 100 million PE reads per sample and paired end 150bp reads (PE 150).

Analysis of sequence quality

Summary statistics regarding the number of raw, trimmed, and non-host reads for each sample were compared using generalized linear models with the "glm" function and the R platform (126) to assess systematic bias across the following sequencing metadata: sequencing run, batch, and lane. For study design metadata, primary comparisons of interest were between CONV vs RWA and typical vs diluted baits. Differences in sequencing results between sample groups were tested with the Wilcoxon signed-rank test when comparing paired values from the same ground beef product (biological replicate) and the Wilcoxon rank-sum test was employed when comparing between treatment groups.

Microbiome and Resistome characterization

To describe the profile of microbes present in ground beef products, reads from 16S rRNA amplicon sequencing were analyzed using the collection of tools contained in Quantitative Insights Into Microbial Ecology version 2 (172). Briefly, all reads are processed for sequence quality and denoising using DADA2 (180), taxonomic classification using a naive bayes classifier trained on the GreenGenes database (97), and removal of chloroplast and mitochondrial

DNA contaminants. Results were then exported into count tables and summarized using the R statistical software (126).

To identify the resistome in ground beef products, the targeted AMR metagenomic sequencing samples were processed using the AMRPlusPlus bioinformatic pipeline and the MEGARes resistance database (49). Further details on the pipeline can be found in the documentation website (http://megares.meglab.org/amrplusplus). Briefly, read trimming and quality filtering is performed using Trimmomatic (122). Host contamination is identified using the Burrows-Wheeler-Aligner (BWA) software (93) with alignment to the Bos Taurus genome (123) and removal of those reads with SamTools (124). These non-host reads were then aligned to the MEGARes database with BWA. Additionally, two updates were added to AMRPlusPlus pipeline; one to deal with PCR duplication introduced with targeted AMR amplification and another to improve the classification of AMR genes requiring the presence of specific SNPs. Read de-duplication was performed with SamTools on the resulting .sam files from alignment to MEGARes.

Only genes accessions with reads aligning to >80% of the reference nucleotide sequence were considered for further analysis, with the exception of reads aligned to genes that require the presence of specific single nucleotide polymorphisms to confer resistance. These reads are identified, extracted from our dataset, and re-classified separately using Resistance Gene Identifier (95) with the “strict” setting which incorporates detection models and CARD's curated similarity cut-offs to increase accuracy in identifying functional AMR genes (see full script: https://github.com/EnriqueDoster/MEG-amrplusplus-RGI-SNPconfirmation). Additionally, in order to investigate the presence of AMR genetic determinants that have been identified as having specific importance to public health when they are identified in human pathogens, this

subset of genes was identified *a priori* and data were searched to identify their presence: (bla(OXA), bla(SME), bla(IMI), bla(NDM), bla(GES), bla(KPC), bla(cphA), bla(TEM), bla(SHV), bla(CTX-M), bla(CMY), vga/vat, cfr).

Count matrix processing

Differences in the microbiome and resistome of RWA and conventional ground beef were analyzed using the R statistical programming software. Cumulative sum scaling (CSS) (131) was used to normalize counts and account for differences in sequencing depth. The resistance data was then summarized to the class and mechanisms level to avoid bias at the "gene" level associated with irregular naming criteria for new resistance genes (132). Statistical analysis was accomplished using the R packages "metagenomeSeq" and "vegan" (130, 131). Sparsely represented resistance gene accessions which were identified in fewer than 5% of samples were removed from further analysis based on published recommendations (131). The taxonomic lineage was identified for each taxon in the microbiome and counts were aggregated to these 6 Linnaean taxonomic levels: phylum, class, order, family, genus, and species. In total, there were 6 count matrices for the microbiome, but to reduce the repetitive reporting of results at all levels and because results at lower taxonomic levels are not considered very reliable (133), statistical results for microbiome are presented at the phylum, class and order levels. In total, 8 unique normalized count matrices (i.e., 6 count matrices describing the microbiome and 2 count matrices characterizing the resistome) were analyzed and reported.

Statistical analysis

The richness and Shannon's diversity was calculated for each sample using "vegan" and statistical comparisons were made using the "wilcox.test" function in R. Normalized counts were Hellinger-transformed (135) for ordination using the metaMDS function from "vegan", which employs non-metric multidimensional scaling on Euclidian distances. Analysis of similarities (ANOSIM) (136) was used to test differences in the microbiome and resistome between treatment groups. Alternatively, to identify which specific features had significantly different numbers of alignments between treatment groups, metagenomeSeq's "fitZig" function was used to fit a zero-inflated Gaussian model and compare log2-fold differences (131). Limma's "makeContrast" function (137) were then used for pairwise comparisons and P-values were adjusted for multiple tests using the Benjamini-Hochberg procedure (138). A value of alpha=0.05 was selected as the statistical significance cut-off value. To account for spurious statistically significant differences in low abundance features, only features with an average expression > 1 were considered.

RESULTS

Sequencing results

Sequencing of the 32 samples processed with AMR target-enrichment produced > 1.3 billion paired end reads (mean: 42,591,676, range: 8,544,874 – 67,058,362). Read quality filtering removed on average 3.7% of raw reads from each sample (range: 3.4% - 4.1%), but the majority of reads were removed from each sample after removal of bovine host contamination (mean: 99.43% of filtered reads per sample, range: 96.8% - 99.95%). With 16S rRNA amplicon sequencing, > 3.1 billion paired end reads were produced (mean: 194,408 reads per sample, range: 100,939 – 219,822). Filtering to improve overall read quality removed on average 7.3% of raw reads from each sample (range: 4.7% - 12.38%). There was a statistical difference in this

number of raw reads produced between label type (P-value < 0.05), but this was likely influenced by a decreased count of reads in samples from vacuum sealed packing type (P-value < 0.05). On average, samples from chub wrap packaging had 51.3 million reads per sample compared to 36.9 million reads per sample in vacuum packaging (Supplemental Table 6.2). Alternatively, there was no statistical difference in the number of 16S amplicon sequencing reads between the treatment groups or packaging types (Supplemental Table 6.3).

Resistome results

Following alignment of reads to the MEGARes AMR database, removal of duplicate reads, and re-classification of reads aligning to gene accessions requiring SNP confirmation with RGI, a total of 267,922 alignments to AMR gene accessions ('hits') were identified across all samples (mean: 8,372 per sample, range: 80 – 51,868). These reads were classified as hits to 565 different gene accessions, which represented genes that confer resistance to 17 different drug classes through 32 distinct resistance mechanisms.

Overall, 87.5% of the classified resistome reads was comprised of hits to tetracycline resistance genetic determinants. Of these hits to genes that confer tetracycline resistance, a majority were classified as encoding tetracycline resistance ribosomal protection proteins and major facilitator superfamily (MFS) efflux pumps (60% and 39%, respectively). The remainder of the resistome sequences was comprised of hits to drug classes in low abundance: 5.3% were hits to multidrug resistance mechanisms (e.g., multidrug efflux pumps), macrolide-lincosamide-streptogramin (3.4%), betalactam (1.8%), aminoglycoside (1.4%), and all other 11 drug classes each making up less than 1% of hits. This pattern of relative abundance for resistome composition was generally consistent across samples (Figure 6.1). Of the list of genes identified

a priori as being important to medicine and public health, bla(CTX-M), bla(CMY), vga/vat, bla(OXA), bla(SME), bla(IMI), bla(cphA), bla(TEM). Overall, these genes were sparsely represented and in total accounted for only 3,439 hits across all 32 samples with bla(TEM) and bla(CTX) genes making up more than half of those counts.

There were no statistically significant differences in the resistome composition between biological replicates (Figure 6.2). Technical replicates were included in all subsequent comparisons between label type. Further, there was no statistically significant difference in the total number of AMR alignments, richness, or Shannon's diversity with either the typical bait processing or half diluted.

ANOSIM testing suggests that the overall resistome composition does not differ between label type at the class level, however there was a statistically significant separation at the mechanism level (ANOSIM R = 0.13, P = 0.018) (Figure 6.3). This difference at the mechanism level appears to be largely influenced by which tetracycline resistance mechanism was dominant in either treatment group. Richness and Shannon's diversity index comparisons were not significantly different between label type at either the class or mechanism levels (Figure 6.4). However, the ZIG model identified that out of 13 drug classes with an average expression > 1, 8 were significantly different between samples with CONV and RWA package labels. The CONV ground beef samples had a significantly higher relative abundance of multi-drug resistance, betalactams, cationic antimicrobial peptides, bacitracin and elfamycin (P-value < 0.05), whereas alignments to trimethoprim and phenicol resistance was more abundant in RWA products (P-value < 0.05**).** While the majority of sample's resistome consisted of tetracycline resistance, the dominant resistance mechanism differs with a significantly higher abundance of Tetracycline

resistance major facilitator superfamily (MFS) efflux pump in CONV samples compared to RWA.

Microbiome results

A total of 585,499 reads were classified taxonomically with an average of 20.15% of reads per sample. In all 2,386 amplicon sequence variants were identified and together represent 27 phyla, 75 classes, and 124 orders. Two phyla, Firmicutes and Proteobacteria, predominated in the microbiome in this study together accounted for > 95% of all normalized counts (65% and 30.2%, respectively) (Figure 6.5). These phyla also consisted of a single taxonomic class making up the majority of counts. The Bacilli class was 97.4% of all Firmicutes alignments and the Gammaproteobacteria class was 85.6% of all Proteobacteria across all samples. Lactobacillales (62.8%, Vibrionales (21.7%), Clostridiales (1.7%), Neisseriales (1.7%), Enterobacteriales (1.7%) and Pseudomonadales (1.2%) were the most abundant taxa at the order level, with the remaining taxa making up less than 1% of classified reads.

Overall, there were no significant differences observed between CONV and RWA microbiomes at the phylum, class, or order level (ANOSIM P-value < 0.05) (Figure 6.6). Similarly, there were no significant differences richness or Shannon's diversity index (Figure 6.7). Indeed, out of 27 phyla, only 3 were significantly different in relative abundance, with CONV samples containing increased proportions of Proteobacteria and decreased Planctomycetes and Chloroflexi compared to RWA samples (P-value < 0.05). Correspondingly, out of 11 classes differentially abundant between treatment groups, Gammaproteobacteria and Clostridia were found in higher relative abundances in CONV samples with the remaining 9 classes found in lower abundance in RWA samples.

While packaging type did not have a significant effect on microbiome composition, sample resistomes significantly separated by source store of purchase (ANOSIM P-value < 0.05) (Figure 6.8). Further, both richness and Shannon's diversity indices were significantly different between CONV and RWA samples at the phylum and class level.

DISCUSSION

Our results suggest that the overall resistome and microbiome in retail ground beef products doesn't differ significantly between ground beef products with CONV and RWA package labels. Consistent with previously reported patterns of AMR in beef production system, the resistome largely consisted of alignments to gene accessions conferring resistance to the drug class, tetracycline (58, 170). Interestingly, while there was no significant separation between resistomes at the AMR mechanism level, CONV samples had a significantly higher relative abundance of the mechanism, tetracycline resistance major facilitator superfamily (MFS) efflux pump.

Though not the goal of this study, our results suggest that the handling of ground beef products by the retail store could potentially influence the microbiome of retail ground beef products. Of note, however, we did not identify significant differences in the resistome between samples from the 6 different stores. Future studies investigating the microbiome in retail ground beef products would benefit from incorporating measures of different product handling protocols by each retail store (e.g. product temperature, time on display, shipping lot number, etc.) to identify the potential drivers of microbiome changes.

In this study we aimed to provide further information about challenges in improving metagenomic sequencing of the bacterial community on ground beef. We provide the first characterization of the microbiome and resistome of individual retail ground beef products (188)

and describe minimal intra-sample variation with resistome sequencing. Based on recommendations from Agilent, we sequenced technical replicates of each sample with half-diluted baits with the goal of reducing host DNA sequencing and improving overall on-target AMR sequencing. Overall, our results suggest that there are no statistically significant differences between technical replicates in sequencing performance or resistome composition.

Metagenomic sequencing is a promising tool for characterizing the microbiome and resistome in retail ground beef products and has potential to be used for tracing individual sequence variants through the food chain. Nonetheless, innovative methods are needed to reduce sequencing cost and improve sequencing depth to get a more detailed perspective of the resistome on ground beef. Furthermore, results from metagenomic sequencing must be carefully considered in parallel with traditional methods to better characterize the risk of AMR in retail products.

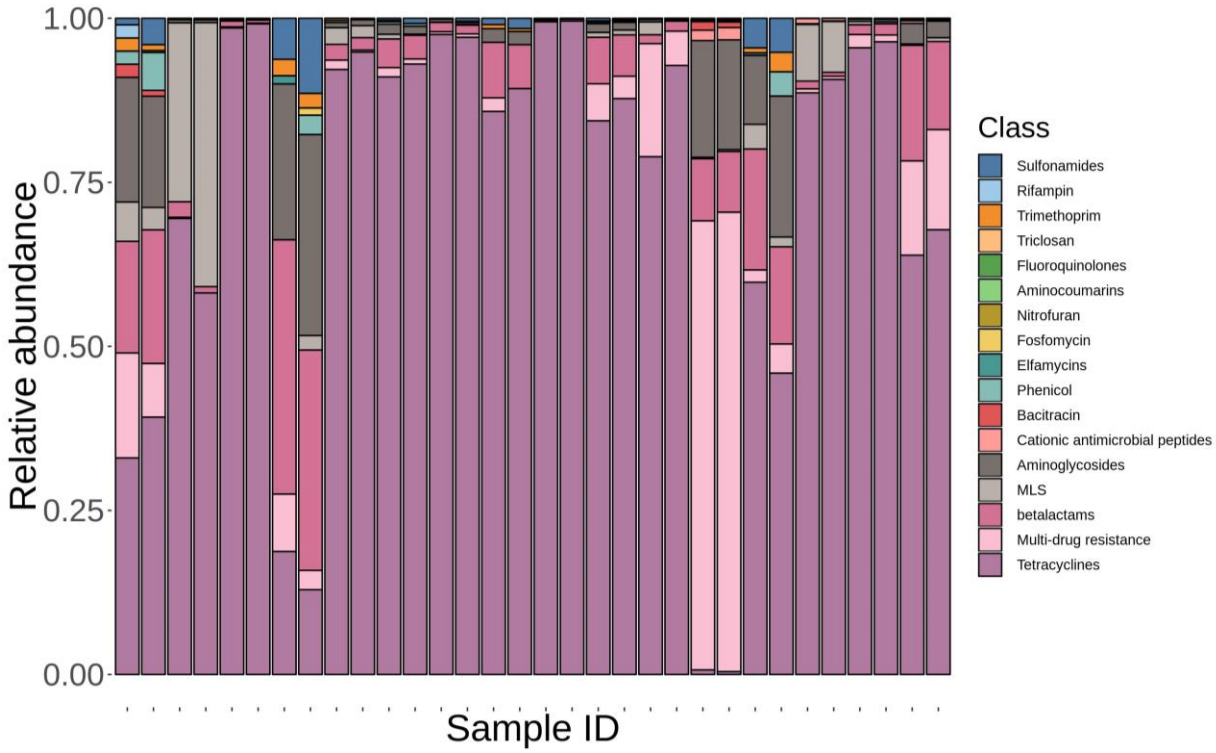


Figure 6.1. Resistome composition for all ground beef samples in the study (N=32) at the drug class level. Individual ground beef samples are on the x-axis and biological replicates are in pairs starting from the left.

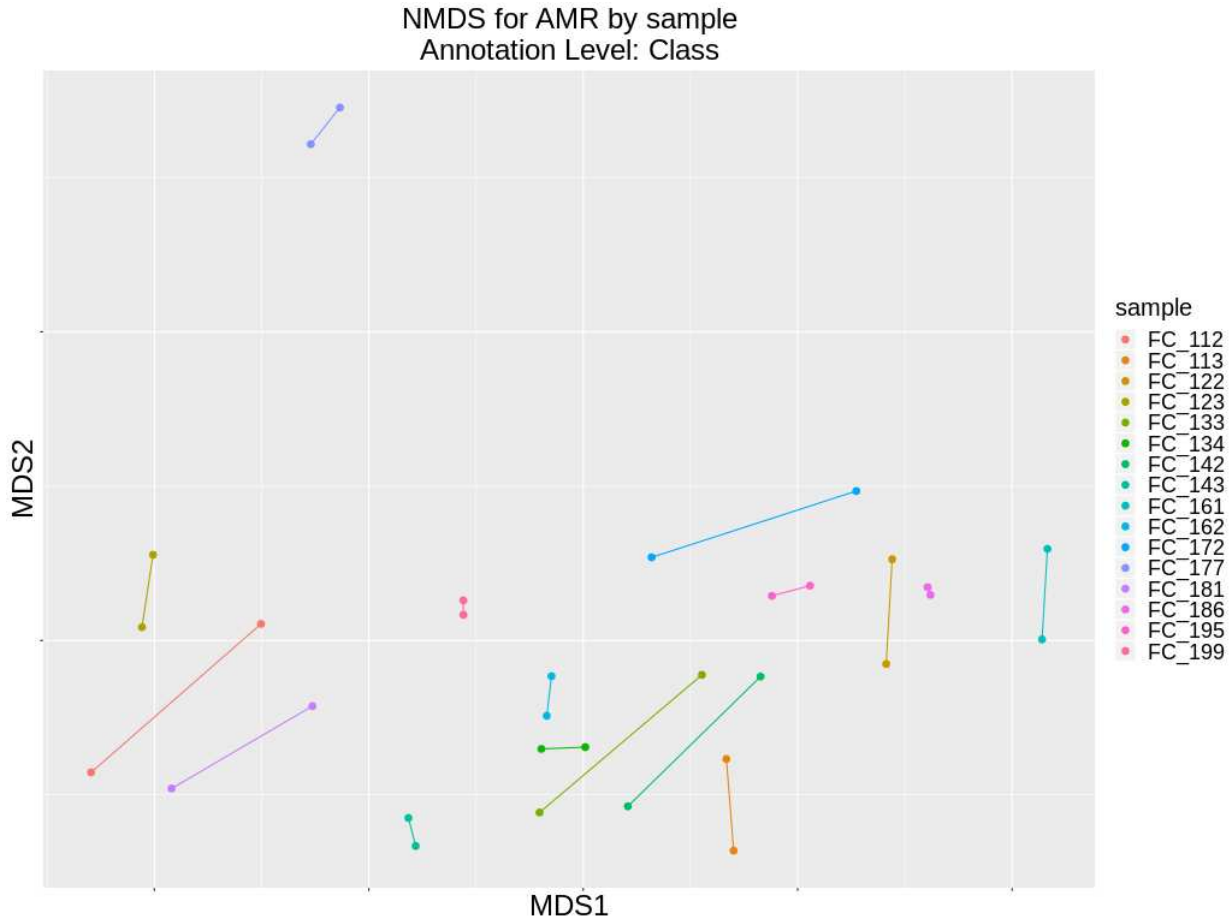


Figure 6.2. Ordination comparing resistome composition at the AMR drug class level, using non-metric multidimensional scaling (NMDS), between biological replicates from the same ground beef product are grouped by color.

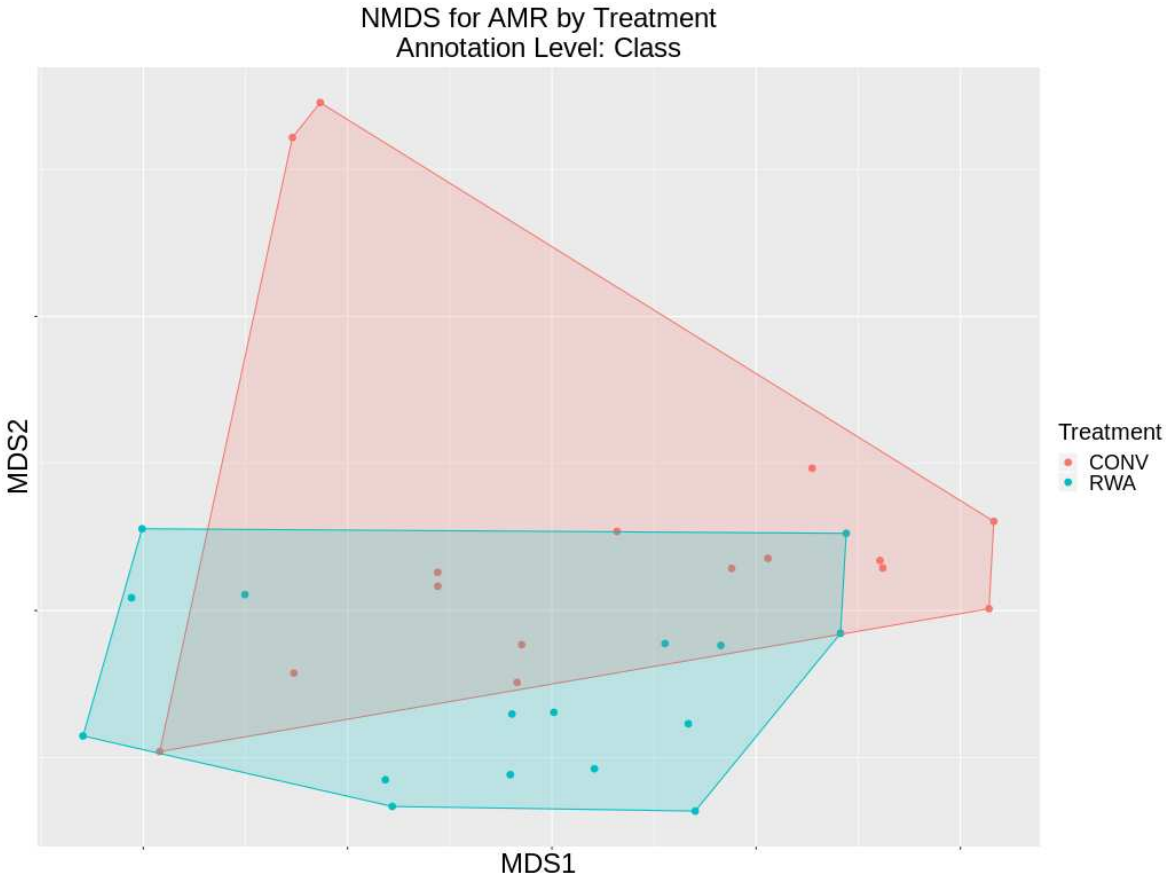


Figure 6.3. Ordination comparing resistome composition at the AMR drug class level, using non-metric multidimensional scaling (NMDS), between labeling types on ground beef products; conventional (CONV) vs raised without antimicrobials (RWA).

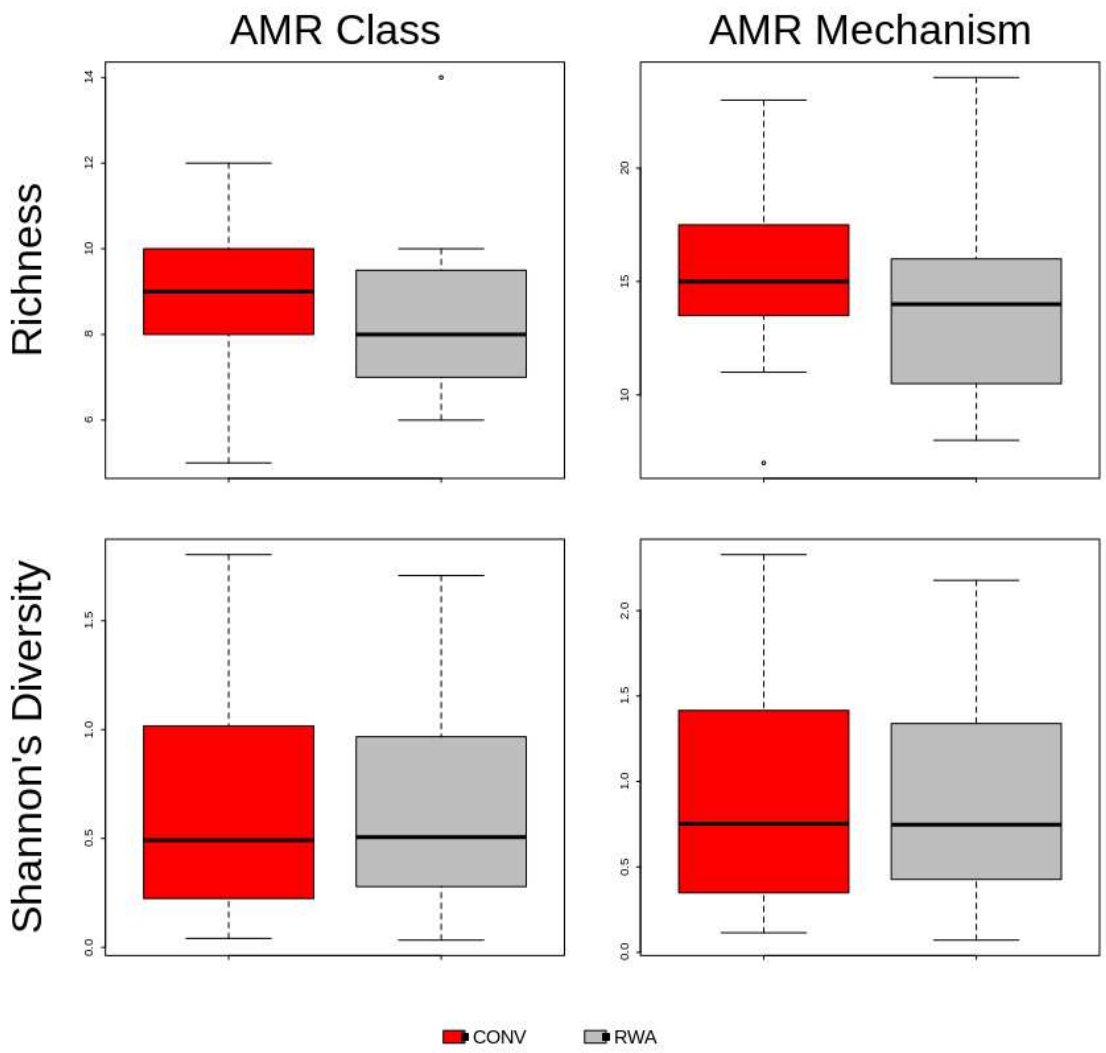


Figure 6.4. Boxplot of resistome richness and Shannon's diversity at the AMR class and mechanism levels of the two study groups, CONV vs RWA. The horizontal line is the median value, the middle box indicates the inter-quantile range, whiskers represent values within 1.5 IQR of the lower and upper quartiles, and individual points show outlier values.

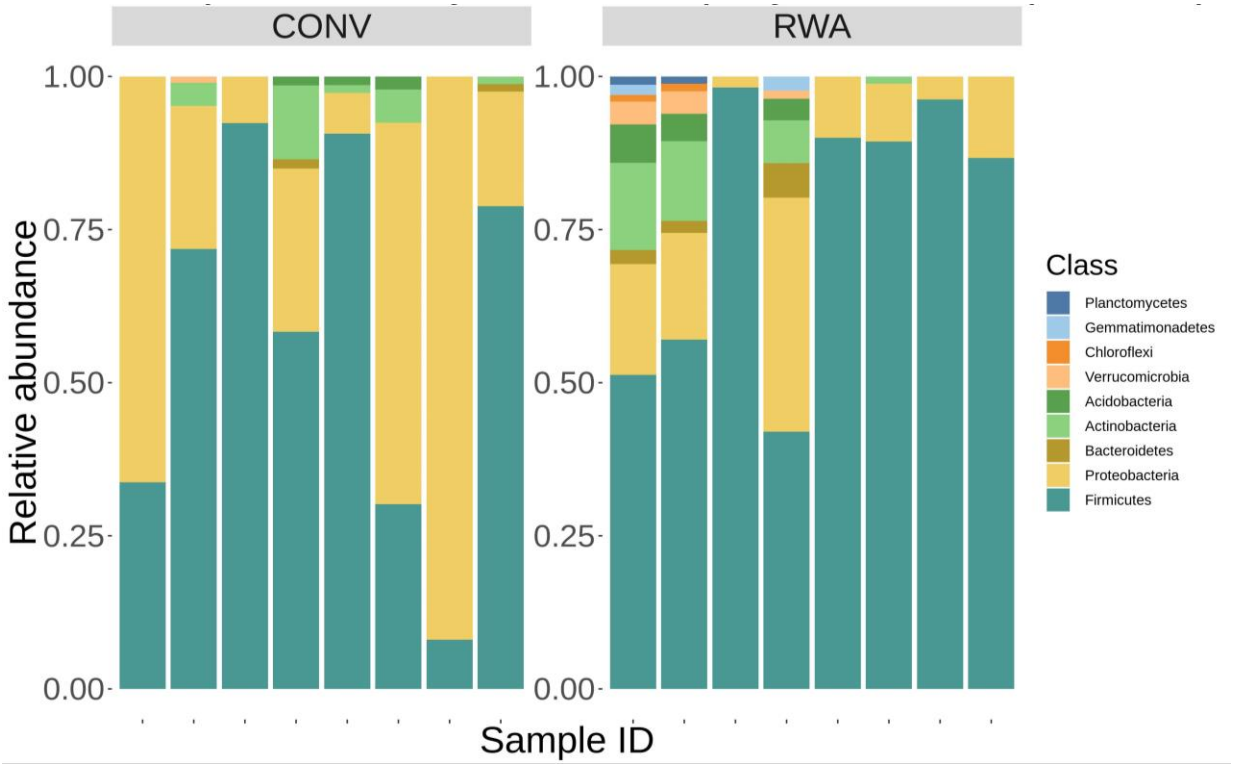


Figure 6.5. Microbiome composition for all ground beef samples in the study (N=32) at the drug class level by labeling type. Individual ground beef samples are on the x-axis and biological replicates are in pairs starting from the left.

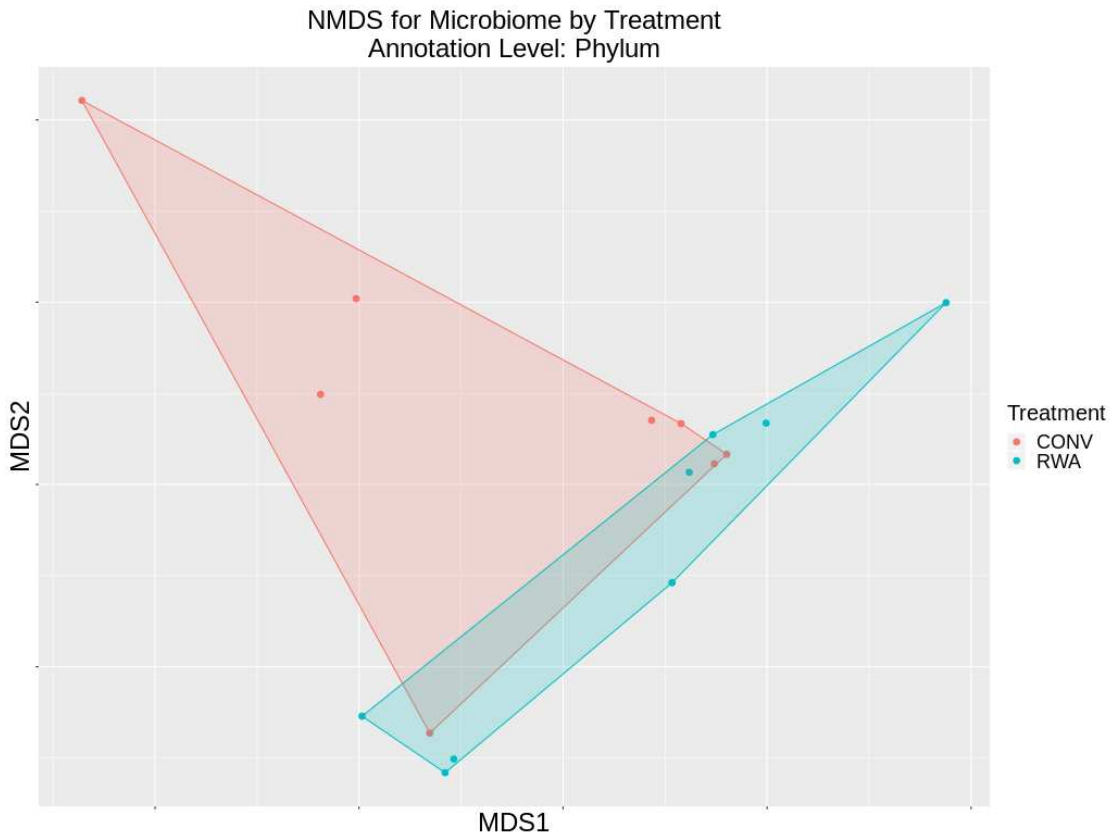


Figure 6.6. Ordination comparing microbiome composition at the phylum taxonomic level, using non-metric multidimensional scaling (NMDS), between labeling types on ground beef products; conventional (CONV) vs raised without antimicrobials (RWA).

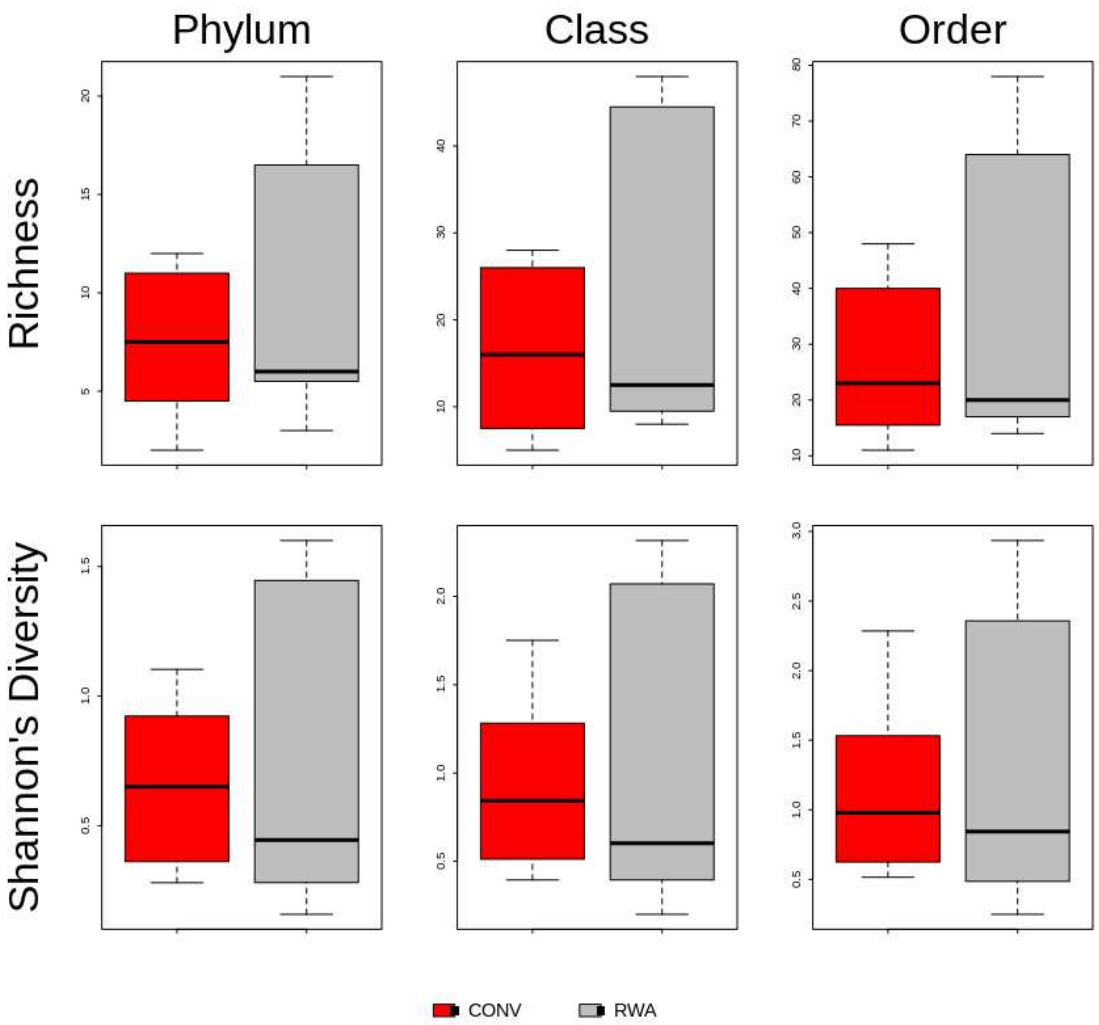


Figure 6.7. Boxplot of resistome richness and Shannon's diversity at the taxonomic phylum, class and order levels of the two study groups, CONV vs RWA. The horizontal line is the median value, the middle box indicates the inter-quantile range, whiskers represent values within 1.5 IQR of the lower and upper quartiles, and individual points show outlier values.

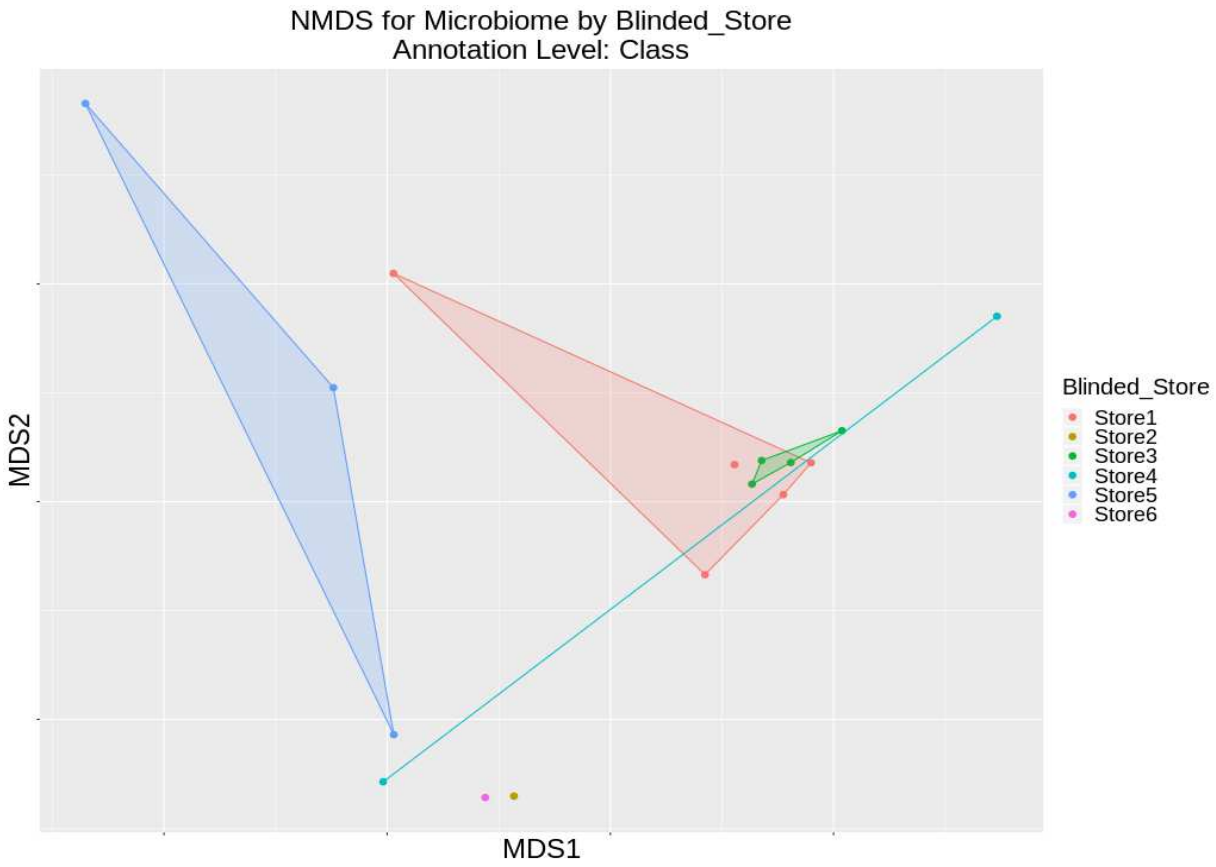


Figure 6.8. Ordination comparing microbiome composition at the phylum taxonomic level, using non-metric multidimensional scaling (NMDS), between the source retail store for each sample.

CONCLUSIONS

The overarching goal of this dissertation was to employ high-throughput sequencing and metagenomics to contribute toward a "real-world" understanding of how antimicrobial AMD use practices in livestock production affects the microbiome and AMR epidemiology by characterizing the resistome. The experiments presented in the preceding dissertation function as a "first look" into different aspects of this question and will hopefully contribute in the work of future research teams.

Generalizing across the results from projects in this dissertation, we presented evidence that the influence of AMD use can be assessed using metagenomics but also suggest that this effect is small in comparison to other factors in the environment. We report that metaphylactic treatment of study cattle did not appear to cause significant shifts in the overall microbiome and resistome by day 11 into the feeding period compared to cattle that did not receive this treatment. Instead, we find that the greatest changes in the microbiome and resistome occur over time in both groups of cattle. Longitudinal studies, preferably starting with animals before shipment, are required to capture the major shifts occurring during transition to the feedlot. We use the same samples from chapter 2 and investigate the utility of metagenomic sequencing for species-level identification of pathogens, in this case Salmonella enterica, by comparing results to aerobic culture, PCR, and lateral flow immunoassay. We reported a lack of concordance between methods and updated our publicly-available bioinformatic pipeline to include steps that improve the relative specificity of species level identification compared to aerobic culture and PCR. Next, we expanded on characterizing how AMD use in beef feedlots effects the microbiome and resistome in beef feedlot cattle and pen floor samples, respectively. We observe similar patterns

with the resistome largely dominated by alignments to genes that confer resistance to the tetracycline drug class and the majority of changes in the lower abundance features associated with time in the feedlot. Finally, we move to the end of the beef production line and explore the microbiome and resistome differences in ground beef with different product labeling types (ie. Conventional vs raised without antimicrobials). There were no major differences in composition between label types or biological replicates from the same sample, and although our study design was only meant to compare between labeling types, our results suggest that the product handling practices at the retail store could have a significant influence on microbiome composition. Overall, our results from constrained analysis suggest we are only explaining a small percentage of the total variance in the microbiome and resistome. There could be many reasons for this, but it is striking that in the 308 samples sequenced metagenomically across the 5 projects in this dissertation, none of the samples reads that could be > 10% classified using current bioinformatic methods. This should sound the alarm that the improvement of reference databases and development of bioinformatic methods, such as probabilistic classification of metagenomic reads, will be crucial for improving the reliability of metagenomic analyses.

We are undergoing a paradigm shift in the field of microbiology broader perspective of microbial ecology is added through metagenomics. In an increasingly complex and global system of food production, agriculture must strive to be at the forefront of adopting new technologies to continuously improve food security and safety for a growing population. However, implementation of new technologies like HTS, metagenomics, and the results garnered from their use must be cautiously interpreted. Importantly, data sharing must be combined with open-source software for bioinformatics and analytic pipelines that wrap software into a user-

friendly package. This is necessary to facilitate collaboration between research teams with distinct fields of study that all benefit from adding a metagenomic perspective to their research.

REFERENCES

1. Abraham EP, Chain E. 1940. An Enzyme from Bacteria able to Destroy Penicillin. Nature 146:837.

2. D'Costa VM, King CE, Kalan L, Morar M, Sung WWL, Schwarz C, Froese D, Zazula G, Calmels F, Debruyne R, Golding GB, Poinar HN, Wright GD. 2011. Antibiotic resistance is ancient. Nature 477:457–461.

3. Davies J, Davies D. 2010. Origins and Evolution of Antibiotic Resistance. Microbiol Mol Biol Rev 74:417–433.

4. Bhullar K, Waglechner N, Pawlowski A, Koteva K, Banks ED, Johnston MD, Barton HA, Wright GD. 2012. Antibiotic Resistance Is Prevalent in an Isolated Cave Microbiome. PLoS ONE 7:e34953.

5. 2013. Antibiotic Resistance Threats in the United States, 2013. The Centers for Disease Control and Prevention.

6. Liu Y-Y, Wang Y, Walsh TR, Yi L-X, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu L-F, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu J-H, Shen J. 2016. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. The Lancet Infectious Diseases 16:161–168.

7. McEwen SA, Collignon PJ. 2018. Antimicrobial Resistance: a One Health Perspective. Microbiol Spectr 6.

8. Spellberg B, Powers JH, Brass EP, Miller LG, Edwards JE. 2004. Trends in antimicrobial drug development: implications for the future. Clin Infect Dis 38:1279–1286.

9. O'Neil J. 2016. Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance. The review on antimicrobial resistance 84.

10. WHO | Antimicrobial resistance: global report on surveillance 2014. WHO.

11. The White House. 2014. National Strategy for Combating Antibiotic-Resistant Bacteria. The White House.

12. World Health Organization. 2011. Critically Important Antimicrobials for Human Medicine, 3rd revision. World Health Organization.

13. Garcia-Alvarez L, Dawson S, Cookson B, Hawkey P. 2012. Working across the veterinary and human health sectors. J Antimicrob Chemother 67 Suppl 1:i37-49.

14. Chang Q, Wang W, Regev-Yochay G, Lipsitch M, Hanage WP. 2015. Antibiotics in agriculture and the risk to human health: how worried should we be? Evol Appl 8:240–247.

15. Modric T, Modric S, Murphy MJ, Bright SJ, Shults S. 2011. Safety of Antibiotic Drugs in Food Animals: Comparison of Findings from Preapproval Studies and Postapproval Experience in the United States with Safety Information in Published Literature. Veterinary Clinics: Food Animal Practice 27:389–405.

16. Marshall BM, Levy SB. 2011. Food Animals and Antimicrobials: Impacts on Human Health. Clin Microbiol Rev 24:718–733.

17. FDA. 2013. Guidance for Industry #213: New Animal Drugs and New Animal Drug Combination Products Administered in or on Medicated Feed or Drinking Water of Food-Producing Animals: Recommendations for Drug Sponsors for Voluntarily Aligning Product Use Conditions with GFI #209. FDA Center for Veterinary Medicine.

18. Li B, Yang Y, Ma L, Ju F, Guo F, Tiedje JM, Zhang T. 2015. Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. ISME J 9:2490–2502.

19. Forsberg KJ, Reyes A, Wang B, Selleck EM, Sommer MOA, Dantas G. 2012. The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. Science 337:1107–1111.

20. Riesenfeld CS, Goodman RM, Handelsman J. 2004. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. Environmental Microbiology 6:981–989.

21. Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J, Handelsman J. 2010. Call of the wild: antibiotic resistance genes in natural environments. Nat Rev Micro 8:251–259.

22. Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) - 2013 Annual Report. Government of Canada, Public Health Agency of Canada.

23. The Danish Integrated Antimicrobial Resistance Monitoring and Research Program. 2013. DANMAP 2013: Use of Antimicrobial Agents and Occurrence of Antimicrobial Resistance

in Bacteria from Food Animals, Foods and Humans in Denmark. The Danish Integrated Antimicrobial Resistance Monitoring and Research Program.

24. SVARM 2014 - Consumption of antibiotics and occurrence of antibiotic resistance in Sweden. Public Health Agency of Sweden and National Veterinary Institute.

25. 2014. Salmonella in U.S. Cattle Feedlots. USDA Animal and Plant Health Inspection Service, Center for Epidemiology and Animal Health, Veterinary Services.

26. Moyaert H, de Jong A, Simjee S, Thomas V. 2014. Antimicrobial resistance monitoring projects for zoonotic and indicator bacteria of animal origin: common aspects and differences between EASSA and EFSA. Vet Microbiol 171:279–283.

27. Boehm AB, Sassoubre LM. 2014. Enterococci as Indicators of Environmental Fecal Contamination, p. . *In* Gilmore, MS, Clewell, DB, Ike, Y, Shankar, N (eds.), Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary, Boston.

28. Alban L, Nielsen EO, Dahl J. 2008. A human health risk assessment for macrolide-resistant Campylobacter associated with the use of macrolides in Danish pig production. Prev Vet Med 83:115–129.

29. Cameron A, McAllister TA. 2016. Antimicrobial usage and resistance in beef production. J Anim Sci Biotechnol 7.

30. Dalhoff A, Ambrose PG, Mouton JW. 2009. A long journey from minimum inhibitory concentration testing to clinically predictive breakpoints: deterministic and probabilistic approaches in deriving breakpoints. Infection 37:296–305.

31. Mouton JW, Brown DFJ, Apfalter P, Cantón R, Giske CG, Ivanova M, MacGowan AP, Rodloff A, Soussy C-J, Steinbakk M, Kahlmeter G. 2012. The role of pharmacokinetics/pharmacodynamics in setting clinical MIC breakpoints: the EUCAST approach. Clin Microbiol Infect 18:E37-45.

32. Stecher B, Denzler R, Maier L, Bernet F, Sanders MJ, Pickard DJ, Barthel M, Westendorf AM, Krogfelt KA, Walker AW, Ackermann M, Dobrindt U, Thomson NR, Hardt W-D. 2012. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. Proc Natl Acad Sci USA 109:1269–1274.

33. Baker KS, Dallman TJ, Field N, Childs T, Mitchell H, Day M, Weill F-X, Lefèvre S, Tourdjman M, Hughes G, Jenkins C, Thomson N. 2018. Horizontal antimicrobial resistance transfer drives epidemics of multiple Shigella species. Nature Communications 9:1462.

34. Benedict KM, Gow SP, Checkley S, Booker CW, McAllister TA, Morley PS. 2013. Methodological comparisons for antimicrobial resistance surveillance in feedlot cattle. BMC Veterinary Research 9:216.

35. Jorgensen JH, Ferraro MJ. 2000. Antimicrobial susceptibility testing: special needs for fastidious organisms and difficult-to-detect resistance mechanisms. Clin Infect Dis 30:799–808.

36. Pulido MR, García-Quintanilla M, Martín-Peña R, Cisneros JM, McConnell MJ. 2013. Progress on the development of rapid methods for antimicrobial susceptibility testing. J Antimicrob Chemother 68:2710–2717.

37. Kircher M, Kelso J. 2010. High-throughput DNA sequencing--concepts and limitations. Bioessays 32:524–536.

38. Mardis ER. 2013. Next-generation sequencing platforms. Annu Rev Anal Chem (Palo Alto Calif) 6:287–303.

39. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. 2012. Performance comparison of benchtop high-throughput sequencing platforms. Nature Biotechnology 30:434–439.

40. Goodwin S, McPherson JD, McCombie WR. 2016. Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics 17:333–351.

41. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13:341.

42. Tripathi R, Sharma P, Chakraborty P, Varadwaj PK. 2016. Next-generation sequencing revolution through big data analytics. Frontiers in Life Science 9:119–149.

43. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, Iyer R, Schatz MC, Sinha S, Robinson GE. 2015. Big Data: Astronomical or Genomical? PLoS Biol 13.

44. Francis SS, Riley LW. 2015. Metagenomic epidemiology: a new frontier. J Epidemiol Community Health 69:306–308.

45. Baquero F. 2012. Metagenomic epidemiology: a public health need for the control of antimicrobial resistance. Clinical Microbiology and Infection 18:67–73.

46. Köser CU, Ellington MJ, Peacock SJ. 2014. Whole-genome sequencing to control antimicrobial resistance. Trends in Genetics 30:401–407.

47. Vartoukian SR, Palmer RM, Wade WG. 2010. Strategies for culture of 'unculturable' bacteria. FEMS Microbiol Lett 309:1–7.

48. Pham VHT, Kim J. 2012. Cultivation of unculturable soil bacteria. Trends in Biotechnology 30:475–484.

49. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, Rovira P, Abdo Z, Jones KL, Ruiz J, Belk KE, Morley PS, Boucher C. 2017. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res 45:D574–D580.

50. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361.

51. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, Jin Q. 2005. VFDB: a reference database for bacterial virulence factors. Nucl Acids Res 33:D325–D328.

52. Pal C, Asiani K, Arya S, Rensing C, Stekel DJ, Larsson DGJ, Hobman JL. 2017. Chapter Seven - Metal Resistance and Its Association With Antibiotic Resistance, p. 261–313. *In* Poole, RK (ed.), Advances in Microbial Physiology. Academic Press.

53. Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ. 2015. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. BMC Genomics 16:964.

54. Herrick JB, Haynes R, Heringa S, Brooks JM, Sobota LT. 2014. Coselection for resistance to multiple late-generation human therapeutic antibiotics encoded on tetracycline resistance plasmids captured from uncultivated stream and soil bacteria. J Appl Microbiol 117:380–389.

55. Holman DB, Timsit E, Booker CW, Alexander TW. 2018. Injectable antimicrobials in commercial feedlot cattle and their effect on the nasopharyngeal microbiota and antimicrobial resistance. Vet Microbiol 214:140–147.

56. Doster E, Rovira P, Noyes NR, Burgess BA, Yang X, Weinroth MD, Lakin SM, Dean CJ, Linke L, Magnuson R, Jones KI, Boucher C, Ruiz J, Belk KE, Morley PS. 2018. Investigating Effects of Tulathromycin Metaphylaxis on the Fecal Resistome and Microbiome of Commercial Feedlot Cattle Early in the Feeding Period. Front Microbiol 9.

57. Thomas M, Webb M, Ghimire S, Blair A, Olson K, Fenske GJ, Fonder AT, Christopher-Hennings J, Brake D, Scaria J. 2017. Metagenomic characterization of the effect of feed additives on the gut microbiome and antibiotic resistome of feedlot cattle. Scientific Reports 7:12257.

58. Noyes NR, Yang X, Linke LM, Magnuson RJ, Dettenwanger A, Cook S, Geornaras I, Woerner DE, Gow SP, McAllister TA, Yang H, Ruiz J, Jones KL, Boucher CA, Morley PS,

Belk KE. 2016. Resistome diversity in cattle and the environment decreases during beef production. eLife 5:e13195.

59. Vikram A, Rovira P, Agga GE, Arthur TM, Bosilevac JM, Wheeler TL, Morley PS, Belk KE, Schmidt JW. 2017. Impact of "Raised without Antibiotics" Beef Cattle Production Practices on Occurrences of Antimicrobial Resistance. Appl Environ Microbiol 83:e01682-17.

60. Chastre J. 2008. Evolving problems with resistant pathogens. Clinical Microbiology and Infection 14:3–14.

61. Mahnert A, Moissl-Eichinger C, Zojer M, Bogumil D, Mizrahi I, Rattei T, Martinez JL, Berg G. 2019. Man-made microbial resistances in built environments. Nature Communications 10:968.

62. Lax S, Sangwan N, Smith D, Larsen P, Handley KM, Richardson M, Guyton K, Krezalek M, Shogan BD, Defazio J, Flemming I, Shakhsheer B, Weber S, Landon E, Garcia-Houchins S, Siegel J, Alverdy J, Knight R, Stephens B, Gilbert JA. 2017. Bacterial colonization and succession in a newly opened hospital. Science Translational Medicine 9:eaah6500.

63. Lax S, Smith DP, Hampton-Marcell J, Owens SM, Handley KM, Scott NM, Gibbons SM, Larsen P, Shogan BD, Weiss S, Metcalf JL, Ursell LK, Vázquez-Baeza Y, Treuren WV, Hasan NA, Gibson MK, Colwell R, Dantas G, Knight R, Gilbert JA. 2014. Longitudinal analysis of microbial interaction between humans and the indoor environment. Science 345:1048–1052.

64. Bergholz TM, Moreno Switt AI, Wiedmann M. 2014. Omics approaches in food safety: fulfilling the promise? Trends in Microbiology 22:275–281.

65. Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, CDC PulseNet Task Force. 2001. PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. Emerging Infect Dis 7:382–389.

66. Tauxe RV. 2006. Molecular subtyping and the transformation of public health. Foodborne Pathog Dis 3:4–8.

67. Gonzalez-Escalona N, Toro M, Rump LV, Cao G, Nagaraja TG, Meng J. 2016. Virulence Gene Profiles and Clonal Relationships of Escherichia coli O26:H11 Isolates from Feedlot Cattle as Determined by Whole-Genome Sequencing. Appl Environ Microbiol 82:3900–3912.

68. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeño-Tárraga AM, Wang H, Holden MT, Wright A, Churcher C, Quail MA, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabbinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B, Parkhill J. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nature Genetics 38:779–786.

69. Fernández-Alarcón C, Singer RS, Johnson TJ. 2011. Comparative Genomics of Multidrug Resistance-Encoding IncA/C Plasmids from Commensal and Pathogenic Escherichia coli from Multiple Animal Sources. PLOS ONE 6:e23415.

70. Delgado-Suárez EJ, Selem-Mojica N, Ortiz-López R, Gebreyes WA, Allard MW, Barona-Gómez F, Rubio-Lozano MS. 2018. Whole genome sequencing reveals widespread

distribution of typhoidal toxin genes and VirB/D4 plasmids in bovine-associated nontyphoidal Salmonella. Sci Rep 8.

71. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Man Kam K, Perez E, Trees E, Kubota K, Takkinen J, Nielsen EM, Carleton H, FWD-NEXT Expert Panel. 2017. PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill 22.

72. Stasiewicz MJ, Oliver HF, Wiedmann M, Bakker HC den. 2015. Whole-Genome Sequencing Allows for Improved Identification of Persistent Listeria monocytogenes in Food-Associated Environments. Appl Environ Microbiol 81:6024–6037.

73. Kwong JC, Mercoulia K, Tomita T, Easton M, Li HY, Bulach DM, Stinear TP, Seemann T, Howden BP. 2016. Prospective Whole-Genome Sequencing Enhances National Surveillance of Listeria monocytogenes. J Clin Microbiol 54:333–342.

74. Schürch AC, van Schaik W. 2017. Challenges and opportunities for whole-genome sequencing-based surveillance of antibiotic resistance. Ann N Y Acad Sci 1388:108–120.

75. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW, Timme R. 2016. Practical Value of Food Pathogen Traceability through Building a Whole-Genome Sequencing Network and Database. J Clin Microbiol 54:1975–1983.

76. Hsieh Y-H, Peterson CM, Raggio A, Keenan MJ, Martin RJ, Ravussin E, Marco ML. 2016. Impact of Different Fecal Processing Methods on Assessments of Bacterial Diversity in the Human Intestine. Front Microbiol 7.

77. Choo JM, Leong LE, Rogers GB. 2015. Sample storage conditions significantly influence faecal microbiome profiles. Scientific Reports 5:16350.

78. Hallmaier-Wacker LK, Lueert S, Roos C, Knauf S. 2018. The impact of storage buffer, DNA extraction method, and polymerase on microbial analysis. Scientific Reports 8:6292.

79. Vo A-TE, Jedlicka JA. 2014. Protocols for metagenomic DNA extraction and Illumina amplicon library preparation for faecal and swab samples. Mol Ecol Resour 14:1183–1197.

80. Henderson G, Cox F, Kittelmann S, Miri VH, Zethof M, Noel SJ, Waghorn GC, Janssen PH. 2013. Effect of DNA Extraction Methods and Sampling Techniques on the Apparent Structure of Cow and Sheep Rumen Microbial Communities. PLoS One 8.

81. van Dijk EL, Jaszczyszyn Y, Thermes C. 2014. Library preparation methods for next-generation sequencing: tone down the bias. Exp Cell Res 322:12–20.

82. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, Ordoukhanian P. 2014. Library construction for next-generation sequencing: Overviews and challenges. Biotechniques 56:61-passim.

83. Feehery GR, Yigit E, Oyola SO, Langhorst BW, Schmidt VT, Stewart FJ, Dimalanta ET, Amaral-Zettler LA, Davis T, Quail MA, Pradhan S. 2013. A Method for Selectively Enriching Microbial DNA from Contaminating Vertebrate Host DNA. PLoS ONE 8:e76096.

84. Noyes NR, Weinroth ME, Parker JK, Dean CJ, Lakin SM, Raymond RA, Rovira P, Doster E, Abdo Z, Martin JN, Jones KL, Ruiz J, Boucher CA, Belk KE, Morley PS. 2017. Enrichment

allows identification of diverse, rare elements in metagenomic resistome-virulome sequencing. Microbiome 5:142.

85. Singer E, Andreopoulos B, Bowers RM, Lee J, Deshpande S, Chiniquy J, Ciobanu D, Klenk H-P, Zane M, Daum C, Clum A, Cheng J-F, Copeland A, Woyke T. 2016. Next generation sequencing data of a defined microbial mock community. Scientific Data 3:sdata201681.

86. Tkacz A, Hortala M, Poole PS. 2018. Absolute quantitation of microbiota abundance in environmental samples. Microbiome 6:110.

87. Deveson IW, Chen WY, Wong T, Hardwick SA, Andersen SB, Nielsen LK, Mattick JS, Mercer TR. 2016. Representing genetic variation with synthetic DNA standards. Nat Methods 13:784–791.

88. Clooney AG, Fouhy F, Sleator RD, Driscoll AO, Stanton C, Cotter PD, Claesson MJ. 2016. Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis. PLOS ONE 11:e0148028.

89. de la Cuesta-Zuluaga J, Escobar JS. 2016. Considerations For Optimizing Microbiome Analysis Using a Marker Gene. Front Nutr 3:26.

90. Lindgreen S, Adair KL, Gardner PP. 2016. An evaluation of the accuracy and speed of metagenome analysis tools. Scientific Reports 6:19233.

91. Hou Y, Wu K, Shi X, Li F, Song L, Wu H, Dean M, Li G, Tsang S, Jiang R, Zhang X, Li B, Liu G, Bedekar N, Lu N, Xie G, Liang H, Chang L, Wang T, Chen J, Li Y, Zhang X, Yang H,

Xu X, Wang L, Wang J. 2015. Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. Gigascience 4.

92. Evans LM, Tahmasbi R, Vrieze SI, Abecasis GR, Das S, Gazal S, Bjelland DW, Candia TR de, Goddard ME, Neale BM, Yang J, Visscher PM, Keller MC. 2018. Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. Nature Genetics 50:737.

93. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997 [q-bio].

94. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, Harris SR. 2017. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microb Genom 3.

95. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, Doshi S, Courtot M, Lo R, Williams LE, Frye JG, Elsayegh T, Sardar D, Westman EL, Pawlowski AC, Johnson TA, Brinkman FSL, Wright GD, McArthur AG. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res 45:D566–D573.

96. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C,

Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44:D733-745.

97. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl Environ Microbiol 72:5069–5072.

98. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Eberhard C, Grüning B, Guerler A, Hillman-Jackson J, Von Kuster G, Rasche E, Soranzo N, Turaga N, Taylor J, Nekrutenko A, Goecks J. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44:W3–W10.

99. Robinson TP, Bu DP, Carrique-Mas J, Fèvre EM, Gilbert M, Grace D, Hay SI, Jiwakanon J, Kakkar M, Kariuki S, Laxminarayan R, Lubroth J, Magnusson U, Thi Ngoc P, Van Boeckel TP, Woolhouse MEJ. 2016. Antibiotic resistance is the quintessential One Health issue. Trans R Soc Trop Med Hyg 110:377–380.

100. Radhouani H, Silva N, Poeta P, Torres C, Correia S, Igrejas G. 2014. Potential impact of antimicrobial resistance in wildlife, environment and human health. Front Microbiol 5.

101. Pallecchi L, Bartoloni A, Paradisi F, Rossolini GM. 2008. Antibiotic resistance in the absence of antimicrobial use: mechanisms and implications. Expert Rev Anti Infect Ther 6:725–732.

102. Bengtsson-Palme J, Kristiansson E, Larsson DGJ. 2018. Environmental factors influencing the development and spread of antibiotic resistance. FEMS Microbiol Rev 42.

103. Bondt N, Jensen VF, Puister-Jansen LF, van Geijlswijk IM. 2013. Comparing antimicrobial exposure based on sales data. Prev Vet Med 108:10–20.

104. P S, S S, S S. 2012. Surveillance and monitoring of antimicrobial resistance and antibiotic consumption in humans and animals 31:105–120.

105. Sanderson MW, Dargatz DA, Wagner BA. 2008. Risk factors for initial respiratory disease in United States' feedlots based on producer-collected daily morbidity counts. Can Vet J 49:373–378.

106. Grissett GP, White BJ, Larson RL. 2015. Structured literature review of responses of cattle to viral and bacterial pathogens causing bovine respiratory disease complex. J Vet Intern Med 29:770–780.

107. Hilton WM. 2014. BRD in 2014: where have we been, where are we now, and where do we want to go? Anim Health Res Rev 15:120–122.

108. Abell KM, Theurer ME, Larson RL, White BJ, Apley M. 2017. A mixed treatment comparison meta-analysis of metaphylaxis treatments for bovine respiratory disease in beef cattle. J Anim Sci 95:626–635.

109. Murray GM, Cassidy JP, Clegg TA, Tratalos JA, McClure J, O'Neill RG, Sammin DJ, Casey MJ, McElroy M, Earley B, Bourke N, More SJ. 2016. A retrospective epidemiological

analysis of risk factors for a primary necropsy diagnosis of bovine respiratory disease. Prev Vet Med 132:49–56.

110. O'Connor AM, Yuan C, Cullen JN, Coetzee JF, da Silva N, Wang C. 2016. A mixed treatment meta-analysis of antibiotic treatment options for bovine respiratory disease - An update. Prev Vet Med 132:130–139.

111. Wellman NG, O'Connor AM. 2007. Meta-analysis of treatment of cattle with bovine respiratory disease with tulathromycin. J Vet Pharmacol Ther 30:234–241.

112. USDA. 2013. Feedlot 2011. Part IV: Health and Health Management on U.S. Feedlots with a Capacity of 1,000 or More Head. USDA Animal and Plant Health Inspection Service.

113. Abdel-Daim MM, Eltaysh R, Hassan A, Mousa SA. 2018. Lycopene Attenuates Tulathromycin and Diclofenac Sodium-Induced Cardiotoxicity in Mice. Int J Mol Sci 19.

114. Er A, Altan F, Cetin G, Dik B, Elmas M, Yazar E. 2011. Assessment of the cardiotoxicity of tulathromycin in rabbits. Acta Vet Hung 59:327–335.

115. Godinho KS. 2008. Susceptibility testing of tulathromycin: Interpretative breakpoints and susceptibility of field isolates. Veterinary Microbiology 129:426–432.

116. Zaheer R, Cook SR, Klima CL, Stanford K, Alexander T, Topp E, Read RR, McAllister TA. 2013. Effect of subtherapeutic vs. therapeutic administration of macrolides on antimicrobial resistance in Mannheimia haemolytica and enterococci isolated from beef cattle. Frontiers in Microbiology 4.

117. Alexander T w., Jin X, Li Q, Cook S, McAllister T a. 2013. Characterization of tetracycline resistance genes in Escherichia coli isolated from feedlot cattle administered therapeutic or subtherapeutic levels of tetracycline. Canadian Journal of Microbiology 59:287–290.

118. Portis E, Lindeman C, Johansen L, Stoltman G. 2012. A ten-year (2000–2009) study of antimicrobial susceptibility of bacteria that cause bovine respiratory disease complex—Mannheimia haemolytica, Pasteurella multocida, and Histophilus somni—in the United States and Canada. J VET Diagn Invest 24:932–944.

119. Pfizer. 2007. Efficacy of DRAXXIN, followed by 7-,10-, or 14-day post-treatment intervals, against naturally occurring bovine respiratory disease in high-risk calves to close. Pfizer Animal Health Technical Bulletin.

120. FDA. 2005. Implantation or Injectable Dosage Form New Animal Drugs; Tulathromycin. Department of Health and Human Services.

121. National Research Council. 2000. Nutrient requirements of beef cattle, 8th revised edition. The National Academic Press.

122. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics btu170.

123. Merchant S, Wood DE, Salzberg SL. 2014. Unexpected cross-species contamination in genome sequencing projects. PeerJ 2.

124. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.

125. Andrews S. 2010. FastQC: A Quality Control tool for High Throughput Sequence Data.

126. Development Core Team R. 2008. R Core Team. R A Language and Environment for Statistical Computing 2014R Foundation for Statistical Computing.

127. Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178–192.

128. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15:R46.

129. Morgulis A, Gertz EM, Schäffer AA, Agarwala R. 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J Comput Biol 13:1028–1040.

130. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2014. vegan: Community Ecology Package.

131. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. Nat Meth 10:1200–1202.

132. Hall RM, Schwarz S. 2016. Resistance gene naming and numbering: is it a new gene or not? J Antimicrob Chemother 71:569–571.

133. Peabody MA, Van Rossum T, Lo R, Brinkman FSL. 2015. Evaluation of shotgun metagenomics sequence classification methods using in silico and in vitro simulated communities. BMC Bioinformatics 16:363.

134. 2016. Tableau (version. 9.1). J Med Libr Assoc 104:182–183.

135. Legendre P, Gallagher ED. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

136. Clarke KR. 1993. Non-parametric multivariate analyses of changes in community structure. Australian Journal of Ecology 18:117–143.

137. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43:e47–e47.

138. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–300.

139. O'Connor AM, Poppe C, McEwen SA. 2002. Changes in the prevalence of resistant Escherichia coli in cattle receiving subcutaneously injectable oxytetracycline in addition to in-feed chlortetracycline compared with cattle receiving only in-feed chlortetracycline. Can J Vet Res 66:145–150.

140. Gaggìa F, Mattarelli P, Biavati B. 2010. Probiotics and prebiotics in animal feeding for safe food production. Int J Food Microbiol 141 Suppl 1:S15-28.

141. McCann JC, Wickersham TA, Loor JJ. 2014. High-throughput Methods Redefine the Rumen Microbiome and Its Relationship with Nutrition and Metabolism. Bioinform Biol Insights 8:109–125.

142. Ghosh S, LaPara TM. 2007. The effects of subtherapeutic antibiotic use in farm animals on the proliferation and persistence of antibiotic resistance among soil bacteria. ISME J 1:191–203.

143. Kyselková M, Jirout J, Vrchotová N, Schmitt H, Elhottová D. 2015. Spread of tetracycline resistance genes at a conventional dairy farm. Front Microbiol 6:536.

144. Platt TM, Loneragan GH, Scott HM, Norby B, Thomson DU, Brown MS, Ives SE, Brashears MM. 2008. Antimicrobial susceptibility of enteric bacteria recovered from feedlot cattle administered chlortetracycline in feed. American Journal of Veterinary Research 69:988–996.

145. Noyes NR, Yang X, Linke LM, Magnuson RJ, Cook SR, Zaheer R, Yang H, Woerner DR, Geornaras I, McArt JA, Gow SP, Ruiz J, Jones KL, Boucher CA, McAllister TA, Belk KE, Morley PS. 2016. Characterization of the resistome in manure, soil and wastewater from dairy and beef production systems. Scientific Reports 6:24645.

146. Holman DB, Timsit E, Alexander TW. 2015. The nasopharyngeal microbiota of feedlot cattle. Scientific Reports 5:15557.

147. Timsit E, Workentine M, Schryvers AB, Holman DB, van der Meer F, Alexander TW. 2016. Evolution of the nasopharyngeal microbiota of beef cattle from weaning to 40 days after arrival at a feedlot. Veterinary Microbiology 187:75–81.

148. Fernando SC, Purvis HT, Najar FZ, Sukharnikov LO, Krehbiel CR, Nagaraja TG, Roe BA, DeSilva U. 2010. Rumen Microbial Population Dynamics during Adaptation to a High-Grain Diet. Appl Environ Microbiol 76:7482–7490.

149. Petri RM, Schwaiger T, Penner GB, Beauchemin KA, Forster RJ, McKinnon JJ, McAllister TA. 2013. Characterization of the Core Rumen Microbiome in Cattle during Transition from Forage to Concentrate as Well as during and after an Acidotic Challenge. PLoS One 8.

150. Yang X, Noyes NR, Doster E, Martin JN, Linke LM, Magnuson RJ, Yang H, Geornaras I, Woerner D, Jones KL, Ruiz J, Boucher C, Morley PS, Belk KE. 2016. Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within their microbiome in beef production chain. Appl Environ Microbiol AEM.00078-16.

151. Parker EPK, Praharaj I, John J, Kaliappan SP, Kampmann B, Kang G, Grassly NC. 2017. Changes in the intestinal microbiota following the administration of azithromycin in a randomised placebo-controlled trial among infants in south India. Scientific Reports 7:9168.

152. Shanks OC, Kelty CA, Archibeque S, Jenkins M, Newton RJ, McLellan SL, Huse SM, Sogin ML. 2011. Community Structures of Fecal Bacteria in Cattle from Different Animal Feeding Operations. Appl Environ Microbiol 77:2992–3001.

153. Chambers L, Yang Y, Littier H, Ray P, Zhang T, Pruden A, Strickland M, Knowlton K. 2015. Metagenomic Analysis of Antibiotic Resistance Genes in Dairy Cow Feces following Therapeutic Administration of Third Generation Cephalosporin. PLOS ONE 10:e0133764.

154. Evans NA. 2005. Tulathromycin: an overview of a new triamilide antibiotic for livestock respiratory disease. Vet Ther 6:83–95.

155. Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. 2017. Shotgun metagenomics, from sampling to analysis. Nature Biotechnology 35:833–844.

156. Hugenholtz P. 2002. Exploring prokaryotic diversity in the genomic era. Genome Biol 3:REVIEWS0003.

157. Staley JT, Konopka A. 1985. Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol 39:321–346.

158. Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, Liachko I, Snelling TJ, Dewhurst RJ, Walker AW, Roehe R, Watson M. 2018. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. Nat Commun 9.

159. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett C, Knight R, Gordon JI. 2007. The human microbiome project: exploring the microbial part of ourselves in a changing world. Nature 449:804–810.

160. Zhou Y, Wylie KM, Feghaly REE, Mihindukulasuriya KA, Elward A, Haslam DB, Storch GA, Weinstock GM. 2016. Metagenomic Approach for Identification of the Pathogens Associated with Diarrhea in Stool Specimens. Journal of Clinical Microbiology 54:368–375.

161. Gire SK, Goba A, Andersen KG, Sealfon RSG, Park DJ, Kanneh L, Jalloh S, Momoh M, Fullah M, Dudas G, Wohl S, Moses LM, Yozwiak NL, Winnicki S, Matranga CB, Malboeuf CM, Qu J, Gladden AD, Schaffner SF, Yang X, Jiang P-P, Nekoui M, Colubri A, Coomber

MR, Fonnie M, Moigboi A, Gbakie M, Kamara FK, Tucker V, Konuwa E, Saffa S, Sellu J, Jalloh AA, Kovoma A, Koninga J, Mustapha I, Kargbo K, Foday M, Yillah M, Kanneh F, Robert W, Massally JLB, Chapman SB, Bochicchio J, Murphy C, Nusbaum C, Young S, Birren BW, Grant DS, Scheiffelin JS, Lander ES, Happi C, Gevao SM, Gnirke A, Rambaut A, Garry RF, Khan SH, Sabeti PC. 2014. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science 345:1369–1372.

162. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A, Hoekstra RM, International Collaboration on Enteric Disease "Burden of Illness" Studies. 2010. The global burden of nontyphoidal Salmonella gastroenteritis. Clin Infect Dis 50:882–889.

163. Laufer AS, Grass J, Holt K, Whichard JM, Griffin PM, Gould LH. 2015. Outbreaks of Salmonella Infections Attributed to Beef – United States, 1973–2011. Epidemiol Infect 143:2003–2013.

164. Burgess BA, Noyes NR, Bolte DS, Hyatt DR, van Metre DC, Morley PS. 2015. Rapid Salmonella detection in experimentally inoculated equine faecal and veterinary hospital environmental samples using commercially available lateral flow immunoassays. Equine Vet J 47:119–122.

165. Burgess BA, Weller CB, Pabilonia KL, Bolte DS, Van Metre DC, Morley PS. 2014. Detection of different serotypes of Salmonella enterica in experimentally inoculated equine fecal samples by commercially available rapid tests. J Vet Intern Med 28:1853–1859.

166. NCBI Resource Coordinators. 2016. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 44:D7-19.

167. Madden T. 2013. The BLAST Sequence Analysis Tool. National Center for Biotechnology Information (US).

168. Miller RR, Montoya V, Gardy JL, Patrick DM, Tang P. 2013. Metagenomics for pathogen detection in public health. Genome Med 5:81.

169. Magouras I, Carmo LP, Stärk KDC, Schüpbach-Regula G. 2017. Antimicrobial Usage and -Resistance in Livestock: Where Should We Focus? Front Vet Sci 4.

170. Benedict KM, Gow SP, McAllister TA, Booker CW, Hannon SJ, Checkley SL, Noyes NR, Morley PS. 2015. Antimicrobial Resistance in Escherichia coli Recovered from Feedlot Cattle and Associations with Antimicrobial Use. PLOS ONE 10:e0143995.

171. Noyes NR, Benedict KM, Gow SP, Waldner CL, Reid-Smith RJ, Booker CW, McALLISTER TA, Morley PS. 2016. Modelling considerations in the analysis of associations between antimicrobial use and resistance in beef feedlot cattle. Epidemiology & Infection 144:1313–1329.

172. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336.

173. Noffsinger T, Lukasiewicz K, Hyder L. 2015. Feedlot Processing and Arrival Cattle Management. Veterinary Clinics: Food Animal Practice 31:323–340.

174. Odeyemi OA, Sani NA. 2016. Antibiotic resistance and burden of foodborne diseases in developing countries. Future Sci OA 2.

175. McEwen SA, Fedorka-Cray PJ. 2002. Antimicrobial Use and Resistance in Animals. Clin Infect Dis 34:S93–S106.

176. Hsueh P-R, Chen W-H, Luh K-T. 2005. Relationships between antimicrobial use and antimicrobial resistance in Gram-negative bacteria causing nosocomial infections from 1991–2003 at a university hospital in Taiwan. International Journal of Antimicrobial Agents 26:463–472.

177. Goossens H, Ferech M, Vander Stichele R, Elseviers M, ESAC Project Group. 2005. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. Lancet 365:579–587.

178. Ashbolt NJ, Amézquita A, Backhaus T, Borriello P, Brandt KK, Collignon P, Coors A, Finley R, Gaze WH, Heberer T, Lawrence JR, Larsson DGJ, McEwen SA, Ryan JJ, Schönfeld J, Silley P, Snape JR, Van den Eede C, Topp E. 2013. Human Health Risk Assessment (HHRA) for Environmental Development and Transfer of Antibiotic Resistance. Environmental Health Perspectives.

179. Hu Y, Gao GF, Zhu B. 2017. The antibiotic resistome: gene flow in environments, animals and human beings. Front Med 11:161–168.

180. Mandal S, Treuren WV, White RA, Eggesbø M, Knight R, Peddada SD. 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. Microbial Ecology in Health and Disease 26:27663.

181. Zaheer R, Noyes N, Polo RO, Cook SR, Marinier E, Domselaar GV, Belk KE, Morley PS, McAllister TA. 2018. Impact of sequencing depth on the characterization of the microbiome and resistome. Scientific Reports 8:5890.

182. Forsberg KJ, Patel S, Gibson MK, Lauber CL, Knight R, Fierer N, Dantas G. 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature 509:612–616.

183. Cytryn E. 2013. The soil resistome: the anthropogenic, the native, and the unknown. Soil Biology & Biochemistry 63:18–23.

184. Wheeler TL, Kalchayanand N, Bosilevac JM. 2014. Pre- and post-harvest interventions to reduce pathogen contamination in the U.S. beef industry. Meat Science 98:372–382.

185. Vikram A, Miller E, Arthur TM, Bosilevac JM, Wheeler TL, Schmidt JW. 2018. Similar Levels of Antimicrobial Resistance in U.S. Food Service Ground Beef Products with and without a “Raised without Antibiotics” Claim. J Food Prot 81:2007–2018.

186. Sajali N, Wong SC, Hanapi UK, Abu Bakar Jamaluddin S, Tasrip NA, Mohd Desa MN. 2018. The Challenges of DNA Extraction in Different Assorted Food Matrices: A Review. J Food Sci 83:2409–2414.

187. McMillin KW. 2017. Advancements in meat packaging. Meat Science 132:153–162.

188. Zhang Y, Boyd SA, Teppen BJ, Tiedje JM, Li H. 2014. Organic acids enhance bioavailability of tetracycline in water to Escherichia coli for uptake and expression of antibiotic resistance. Water Research 65:98–106.

Supplemental Table 2.1. Metadata information for sequencing results for all 60 samples.

Sample ID	Time	Treatment	Number of paired-reads Pre-processing	Number of paired-reads After trimming	Number of paired non-host reads After removing bovine sequences	Number of paired non-host reads classified with kraken	AMR abundance normalized with Li's equation
11_GGCTAC_L001	Day1	Treated	72,985,725	69,992,146	51,643,315	606,695	0.216
37_CTTGTA_L001	Day1	Treated	45,656,078	43,004,261	37,718,769	2,794,500	0.29
51_AGTCAA_L001	Day1	Treated	74,700,329	71,482,218	59,626,525	3,173,217	0.224
62_AGTTCC_L002	Day1	Treated	53,294,398	50,859,060	43,158,901	5,289,938	0.304
83_ATGTCA_L002	Day1	Treated	55,614,995	52,232,070	39,290,183	2,530,506	0.174
84_CCGTCC_L003	Day1	Treated	51,280,133	48,350,438	38,255,539	4,633,854	0.264
92_GTCCGC_L003	Day1	Treated	54,910,092	51,993,321	43,325,943	3,689,555	0.267
93_GTGAAA_L003	Day1	Treated	58,033,031	55,580,928	40,559,750	3,752,319	0.141
102_GTGGCC_L004	Day1	Treated	55,591,833	53,370,571	44,638,067	2,553,827	0.244
103_ATTCCT_L004	Day1	Treated	52,596,429	51,252,192	43,377,011	2,553,209	0.298
130_ATCACG_L005	Day1	Treated	55,555,652	53,615,411	41,097,185	3,494,778	0.204

155_CGATGT_L005	Day1	Treated	76,629,319	73,138,923	58,400,529	1,944,852	0.316
156_ACAGTG_L005	Day1	Treated	28,066,461	26,719,570	21,964,644	1,839,227	0.262
158_GCCAAT_L006	Day1	Treated	53,539,553	50,648,293	45,309,724	2,191,052	0.316
164_TTAGGC_L006	Day1	Treated	41,062,771	39,545,103	31,925,154	3,159,015	0.406
208_CAGATC_L007	Day1	Untreated	26,209,455	24,443,697	20,612,673	3,335,455	0.182
216_ACTTGA_L007	Day1	Untreated	70,016,669	67,415,572	56,255,245	3,498,544	0.26
220_TGACCA_L007	Day1	Untreated	21,907,828	20,695,675	17,656,228	4,102,371	0.216
227_ATGTCA_L008	Day1	Untreated	55,312,466	52,703,369	45,163,638	2,771,254	0.318
228_CCGTCC_L008	Day1	Untreated	34,080,311	32,659,118	28,514,109	2,624,568	0.176
261_GTCCGC_L001	Day1	Untreated	34,338,217	32,563,714	26,198,981	4,182,001	0.203
281_GTGAAA_L001	Day1	Untreated	36,176,709	34,445,214	29,577,392	2,551,960	0.205
285_GTGGCC_L001	Day1	Untreated	63,607,565	61,119,500	46,706,718	1,421,774	0.169
286_GTTTCG_L002	Day1	Untreated	40,564,315	38,467,390	34,000,756	2,913,129	0.228
287_CGTACG_L002	Day1	Untreated	46,786,199	44,139,338	36,280,650	1,463,575	0.225

289_GAGTGG_L003	Day1	Untreated	51,425,202	47,202,248	38,356,312	2,577,947	0.195
298_ACTGAT_L003	Day1	Untreated	54,748,373	51,867,613	44,958,518	2,831,475	0.173
349_TGACCA_L003	Day1	Untreated	30,047,796	28,034,966	24,351,778	1,219,836	0.376
372_CAGATC_L004	Day1	Untreated	23,736,703	22,683,717	16,895,206	2,726,636	0.138
376_ACTTGA_L004	Day1	Untreated	58,771,628	56,205,618	42,557,433	2,714,509	0.177
11p2_GATCAG_L001	Day 11	Treated	29,669,806	28,389,347	23,895,589	3,237,473	0.449
37p2_TAGCTT_L001	Day 11	Treated	52,614,262	50,404,105	41,138,842	3,085,320	0.506
51p2_GGCTAC_L002	Day 11	Treated	46,571,417	44,048,799	38,220,147	1,647,468	0.524
62p2_CTTGTA_L002	Day 11	Treated	47,141,374	44,778,876	37,172,830	3,119,388	0.453
83p2_AGTCAA_L002	Day 11	Treated	65,548,349	62,394,918	39,853,404	2,370,012	0.305
84p2_AGTTCC_L003	Day 11	Treated	34,443,438	32,743,305	26,206,629	3,744,889	0.623
92p2_ATCACG_L003	Day 11	Treated	66,782,066	64,222,092	51,185,393	1,158,742	0.266

93p2_TTAGGC_L004	Day 11	Treated	48,348,096	47,031,397	37,799,483	2,716,343	0.408
102p2_TGACCA_L004	Day 11	Treated	30,670,200	29,448,145	23,908,771	2,369,338	0.146
103p2_ACTGAT_L004	Day 11	Treated	38,074,219	37,083,085	31,243,539	2,219,401	0.328
130p2_ATTCCT_L005	Day 11	Treated	61,997,511	60,214,244	48,740,862	4,057,605	0.288
155p2_GTGGCC_L005	Day 11	Treated	29,739,401	28,764,726	18,320,679	2,928,288	0.109
156p2_GTTTCG_L006	Day 11	Treated	49,975,590	48,532,827	35,376,729	2,290,594	0.352
158p2_CGTACG_L006	Day 11	Treated	80,355,147	76,323,857	68,065,491	3,965,980	0.398
164p2_GAGTGG_L006	Day 11	Treated	26,823,553	25,322,716	22,311,700	3,096,126	0.191
208p2_TAGCTT_L007	Day 11	Untreated	61,286,424	59,749,182	48,543,446	3,268,122	0.377
216p2_GGCTAC_L007	Day 11	Untreated	42,686,592	41,214,734	36,314,805	1,421,234	0.488
220p2_CTTGTA_L008	Day 11	Untreated	13,489,760	12,754,079	9,468,742	5,836,885	0.179
227p2_AGTCAA_L008	Day 11	Untreated	54,923,602	53,563,269	41,505,387	3,322,451	0.281

228p2_AGTTCC_L008	Day 11	Untreated	55,020,213	53,474,750	40,519,986	3,235,590	0.288
261p2_ATGTCA_L001	Day 11	Untreated	54,771,620	52,546,916	39,561,059	3,360,476	0.189
281p2_CCGTCC_L001	Day 11	Untreated	47,987,247	46,045,972	37,891,591	2,347,754	0.317
285p2_GTGAAA_L002	Day 11	Untreated	71,610,406	68,824,576	54,010,235	5,102,863	0.328
286p2_GTCCGC_L002	Day 11	Untreated	45,363,197	43,373,764	32,567,904	2,872,877	0.233
287p2_ATCACG_L002	Day 11	Untreated	56,833,758	54,532,170	40,870,571	1,727,786	0.479
289p2_CGATGT_L003	Day 11	Untreated	70,450,534	66,767,595	43,834,559	1,646,611	0.426
298p2_TTAGGC_L003	Day 11	Untreated	50,745,968	48,555,641	35,772,337	3,786,022	0.336
349p2_CGATGT_L004	Day 11	Untreated	58,939,701	56,391,886	44,845,183	3,440,191	0.207
372p2_ACAGTG_L004	Day 11	Untreated	35,654,430	33,845,122	30,347,544	1,676,403	0.484
376p2_GCCAAT_L004	Day 11	Untreated	40,178,825	38,453,735	33,873,573	3,109,820	0.476

Class	Mechanism	Day1-Treated	Day1-Untreated	Day11-Treated	Day11-Untreated
Tetracyclines	Tetracycline resistance ribosomal protection proteins	15	15	15	15
	Tetracycline resistance major facilitator superfamily MFS efflux pumps	15	14	9	14
	Tetracycline inactivation enzymes	4	6	3	9
MLS	Macrolide resistance efflux pumps	15	15	15	15
	Lincosamide nucleotidyltransferases	15	15	15	15
	23S rRNA methyltransferases	15	15	15	15
	Macrolide phosphotransferases	1			8
betalactams	Class A betalactamases	15	15	15	15
	Penicillin binding protein	10	7	4	8
	Class C betalactamases	11	8	3	6
Aminoglycosides	Aminoglycoside O-nucleotidyltransferases	14	13	15	15
	Aminoglycoside O-phosphotransferases	8	3	15	15
	Aminoglycoside efflux pumps	12	10	4	8
	Aminoglycoside N-acetyltransferases	1	1	11	13
	Aminoglycoside efflux regulator	10	8	6	7
Multi-drug resistance	Multi-drug efflux pumps	11	11	5	10
	MDR regulator	12	11	5	12
	MDR mutant porin proteins	4	2		1
Cationic antimicrobial peptides	Lipid A modification	11	5	3	8
	Polymyxin B resistance regulator	11	10	3	7
Phenicol	Chloramphenicol acetyltransferases			7	8
	Phenicol 23S rRNA methyltransferases			1	2
Fluoroquinolones	Fluoroquinolone-resistant DNA topoisomerases	4	3	1	2
Bacitracin	Undecaprenyl pyrophosphate phosphatase	7	3	1	2

Supplemental Figure 2.1. Heatmap showing the number of samples in each group (n=15) along the columns with alignments to AMR mechanisms in the rows.

Supplemental Table 3.1. Culture, PCR, and sequencing results for all 60 samples.

Sample name	Time	Treatment	PCR	Culture Serogroup	[DNA] ng/ul	260:280	Yield (Mbases)	Nonhost paired reads
11_GGCTAC_L001	Arrival	Treated	0	0	29.5	1.54	18392	51643315
37_CTTGTA_L001	Arrival	Treated	0	0	49.8	1.49	11505	37718769
51_AGTCAA_L001	Arrival	Treated	0	0	30.8	1.69	18824	59626525
62_AGTTCC_L002	Arrival	Treated	0	0	30.4	1.59	13430	43158901
83_ATGTCA_L002	Arrival	Treated	0	0	28.3	1.62	14015	39290183
84_CCGTCC_L003	Arrival	Treated	0	0	22.3	1.57	12923	38255539
92_GTCCGC_L003	Arrival	Treated	0	0	22.9	1.77	13837	43325943
93_GTGAAA_L003	Arrival	Treated	0	0	34.7	1.5	14624	40559750
102_GTGGCC_L004	Arrival	Treated	0	0	39.2	1.52	14009	44638067
103_ATTCCT_L004	Arrival	Treated	0	0	51.7	1.66	13254	43377011
130_ATCACG_L005	Arrival	Treated	0	0	21.7	1.51	14000	41097185
155_CGATGT_L005	Arrival	Treated	1	K	26.7	1.63	19311	58400529
156_ACAGTG_L005	Arrival	Treated	0	0	28.6	1.49	7073	21964644
158_GCCAAT_L006	Arrival	Treated	0	0	87.7	1.51	13492	45309724
164_TTAGGC_L006	Arrival	Treated	1	C1	25	1.58	10348	31925154
208_CAGATC_L007	Arrival	Untreated	0	0	57.6	1.47	6605	20612673
216_ACTTGA_L007	Arrival	Untreated	0	0	21.4	1.56	17644	56255245
220_TGACCA_L007	Arrival	Untreated	0	0	58.6	1.74	5521	17656228
227_ATGTCA_L008	Arrival	Untreated	0	0	21	1.47	13939	45163638

228_CCGTCC_L008	Arrival	Untreated	0	0	30.3	1.42	8588	28514109
261_GTCCGC_L001	Arrival	Untreated	0	0	25.9	1.49	8653	26198981
281_GTGAAA_L001	Arrival	Untreated	0	0	72.7	1.61	9117	29577392
285_GTGGCC_L001	Arrival	Untreated	0	0	25	1.51	16029	46706718
286_GTTTCG_L002	Arrival	Untreated	0	0	95.3	1.36	10222	34000756
287_CGTACG_L002	Arrival	Untreated	0	0	57.4	1.53	11790	36280650
289_GAGTGG_L003	Arrival	Untreated	0	0	70.4	1.51	12959	38356312
298_ACTGAT_L003	Arrival	Untreated	0	0	50.1	1.41	13797	44958518
349_TGACCA_L003	Arrival	Untreated	0	0	53.5	1.59	7572	24351778
372_CAGATC_L004	Arrival	Untreated	0	0	22.5	1.56	5982	16895206
376_ACTTGA_L004	Arrival	Untreated	0	0	26.5	1.5	14810	42557433
11p2_GATCAG_L001	Day11	Treated	0	0	91.9	1.58	7477	23895589
37p2_TAGCTT_L001	Day11	Treated	0	0	44.7	1.7	13259	41138842
51p2_GGCTAC_L002	Day11	Treated	0	0	69.9	1.63	11736	38220147
62p2_CTTGTA_L002	Day11	Treated	0	0	23.9	1.82	11880	37172830
83p2_AGTCAA_L002	Day11	Treated	0	0	32.8	1.67	16518	39853404
84p2_AGTTCC_L003	Day11	Treated	0	0	27.6	1.81	8680	26206629
92p2_ATCACG_L003	Day11	Treated	0	0	21.4	1.73	16829	51185393
93p2_TTAGGC_L004	Day11	Treated	0	0	39.8	1.65	12184	37799483
102p2_TGACCA_L004	Day11	Treated	0	0	55.8	1.69	7729	23908771
103p2_ACTGAT_L004	Day11	Treated	0	0	65.8	1.73	9595	31243539
130p2_ATTCCT_L005	Day11	Treated	0	0	64.5	1.76	15623	48740862

155p2_GTGGCC_L005	Day11	Treated	0	0	21.5	1.83	7494	18320679
156p2_GTTTCG_L006	Day11	Treated	1	C1	23.7	1.93	12594	35376729
158p2_CGTACG_L006	Day11	Treated	0	0	57.4	1.5	20249	68065491
164p2_GAGTGG_L006	Day11	Treated	0	0	84.8	1.63	6760	22311700
208p2_TAGCTT_L007	Day11	Untreated	0	0	62.7	1.65	15444	48543446
216p2_GGCTAC_L007	Day11	Untreated	0	0	57.9	1.69	10757	36314805
220p2_CTTGTA_L008	Day11	Untreated	0	0	21.1	1.84	3399	9468742
227p2_AGTCAA_L008	Day11	Untreated	0	0	69.4	1.69	13841	41505387
228p2_AGTTCC_L008	Day11	Untreated	0	0	48.7	1.67	13865	40519986
261p2_ATGTCA_L001	Day11	Untreated	0	0	54.7	1.62	13802	39561059
281p2_CCGTCC_L001	Day11	Untreated	0	0	43.1	1.73	12093	37891591
285p2_GTGAAA_L002	Day11	Untreated	0	0	25.8	1.63	18046	54010235
286p2_GTCCGC_L002	Day11	Untreated	0	0	28.4	1.69	11432	32567904
287p2_ATCACG_L002	Day11	Untreated	0	0	25.5	1.87	14322	40870571
289p2_CGATGT_L003	Day11	Untreated	0	0	54	1.72	17754	43834559
298p2_TTAGGC_L003	Day11	Untreated	0	0	19.1	1.52	12788	35772337
349p2_CGATGT_L004	Day11	Untreated	0	0	57.8	1.78	14853	44845183
372p2_ACAGTG_L004	Day11	Untreated	0	0	21.9	1.69	8985	30347544
376p2_GCCAAT_L004	Day11	Untreated	0	0	31.6	1.7	10125	33873573

Supplemental Table 3.2. Metagenomic results for *Salmonella enterica* identification using 4 different kraken databases across 60 samples.

Species name	Sample ID	Raw_mapped_reads	Type
Salmonella enterica	102_GTGGCC_L004	2766	Salmonella_custom_v2
Salmonella enterica	102_GTGGCC_L004	3231	Salmonella_standard_v2
Salmonella enterica	102p2_TGACCA_L004	1329	Salmonella_custom_v2
Salmonella enterica	102p2_TGACCA_L004	1440	Salmonella_standard_v2
Salmonella enterica	103_ATTCCT_L004	5	Salmonella_custom_filter_v2
Salmonella enterica	103_ATTCCT_L004	5	Salmonella_custom_filter_v2_ misclassified Salmonella
Salmonella enterica	103_ATTCCT_L004	4783	Salmonella_custom_v2
Salmonella enterica	103_ATTCCT_L004	5	Salmonella_standard_filter_v2
Salmonella enterica	103_ATTCCT_L004	6029	Salmonella_standard_v2
Salmonella enterica	103p2_ACTGAT_L004	1827	Salmonella_custom_v2
Salmonella enterica	103p2_ACTGAT_L004	2084	Salmonella_standard_v2
Salmonella enterica	11_GGCTAC_L001	2414	Salmonella_custom_v2
Salmonella enterica	11_GGCTAC_L001	2818	Salmonella_standard_v2
Salmonella enterica	11p2_GATCAG_L001	1060	Salmonella_custom_v2

Salmonella enterica	11p2_GATCAG_L001	1325	Salmonella_standard_v2
Salmonella enterica	130_ATCACG_L005	2071	Salmonella_custom_v2
Salmonella enterica	130_ATCACG_L005	2521	Salmonella_standard_v2
Salmonella enterica	130p2_ATTCCT_L005	2144	Salmonella_custom_v2
Salmonella enterica	130p2_ATTCCT_L005	2590	Salmonella_standard_v2
Salmonella enterica	155_CGATGT_L005	3734	Salmonella_custom_v2
Salmonella enterica	155_CGATGT_L005	4471	Salmonella_standard_v2
Salmonella enterica	155p2_GTGGCC_L005	856	Salmonella_custom_v2
Salmonella enterica	155p2_GTGGCC_L005	1034	Salmonella_standard_v2
Salmonella enterica	156_ACAGTG_L005	1	Salmonella_custom_filter_v2
Salmonella enterica	156_ACAGTG_L005	1	Salmonella_custom_filter_v2_ misclassified Salmonella
Salmonella enterica	156_ACAGTG_L005	1341	Salmonella_custom_v2
Salmonella enterica	156_ACAGTG_L005	1	Salmonella_standard_filter_v2
Salmonella enterica	156_ACAGTG_L005	1536	Salmonella_standard_v2
Salmonella enterica	156p2_GTTTCG_L006	1625	Salmonella_custom_v2
Salmonella enterica	156p2_GTTTCG_L006	2129	Salmonella_standard_v2

Salmonella enterica	158_GCCAAT_L006	2	Salmonella_custom_filter_v2
Salmonella enterica	158_GCCAAT_L006	2	Salmonella_custom_filter_v2_ misclassified Salmonella
Salmonella enterica	158_GCCAAT_L006	2499	Salmonella_custom_v2
Salmonella enterica	158_GCCAAT_L006	3	Salmonella_standard_filter_v2
Salmonella enterica	158_GCCAAT_L006	2940	Salmonella_standard_v2
Salmonella enterica	158p2_CGTACG_L006	3206	Salmonella_custom_v2
Salmonella enterica	158p2_CGTACG_L006	3654	Salmonella_standard_v2
Salmonella enterica	164_TTAGGC_L006	3	Salmonella_custom_filter_v2
Salmonella enterica	164_TTAGGC_L006	3	Salmonella_custom_filter_v2_ misclassified Salmonella
Salmonella enterica	164_TTAGGC_L006	1829	Salmonella_custom_v2
Salmonella enterica	164_TTAGGC_L006	3	Salmonella_standard_filter_v2
Salmonella enterica	164_TTAGGC_L006	2224	Salmonella_standard_v2
Salmonella enterica	164p2_GAGTGG_L006	1084	Salmonella_custom_v2
Salmonella enterica	164p2_GAGTGG_L006	1234	Salmonella_standard_v2
Salmonella enterica	208_CAGATC_L007	1096	Salmonella_custom_v2
Salmonella enterica	208_CAGATC_L007	1261	Salmonella_standard_v2

Salmonella enterica	208p2_TAGCTT_L007	2244	Salmonella_custom_v2
Salmonella enterica	208p2_TAGCTT_L007	2787	Salmonella_standard_v2
Salmonella enterica	216_ACTTGA_L007	3082	Salmonella_custom_v2
Salmonella enterica	216_ACTTGA_L007	3534	Salmonella_standard_v2
Salmonella enterica	216p2_GGCTAC_L007	1726	Salmonella_custom_v2
Salmonella enterica	216p2_GGCTAC_L007	2049	Salmonella_standard_v2
Salmonella enterica	220_TGACCA_L007	960	Salmonella_custom_v2
Salmonella enterica	220_TGACCA_L007	1118	Salmonella_standard_v2
Salmonella enterica	220p2_CTTGTA_L008	392	Salmonella_custom_v2
Salmonella enterica	220p2_CTTGTA_L008	424	Salmonella_standard_v2
Salmonella enterica	227_ATGTCA_L008	12	Salmonella_custom_filter_v2
Salmonella enterica	227_ATGTCA_L008	12	Salmonella_custom_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	227_ATGTCA_L008	2797	Salmonella_custom_v2
Salmonella enterica	227_ATGTCA_L008	12	Salmonella_standard_filter_v2
Salmonella enterica	227_ATGTCA_L008	12	Salmonella_standard_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	227_ATGTCA_L008	3266	Salmonella_standard_v2

Salmonella enterica	227p2_AGTCAA_L008	2441	Salmonella_custom_v2
Salmonella enterica	227p2_AGTCAA_L008	2703	Salmonella_standard_v2
Salmonella enterica	228_CCGTCC_L008	1521	Salmonella_custom_v2
Salmonella enterica	228_CCGTCC_L008	1	Salmonella_standard_filter_v2
Salmonella enterica	228_CCGTCC_L008	1	Salmonella_standard_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	228_CCGTCC_L008	1785	Salmonella_standard_v2
Salmonella enterica	228p2_AGTTCC_L008	2214	Salmonella_custom_v2
Salmonella enterica	228p2_AGTTCC_L008	1	Salmonella_standard_filter_v2
Salmonella enterica	228p2_AGTTCC_L008	1	Salmonella_standard_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	228p2_AGTTCC_L008	2556	Salmonella_standard_v2
Salmonella enterica	261_GTCCGC_L001	1409	Salmonella_custom_v2
Salmonella enterica	261_GTCCGC_L001	3	Salmonella_standard_filter_v2
Salmonella enterica	261_GTCCGC_L001	1687	Salmonella_standard_v2
Salmonella enterica	261p2_ATGTCA_L001	2241	Salmonella_custom_v2
Salmonella enterica	261p2_ATGTCA_L001	2645	Salmonella_standard_v2
Salmonella enterica	281_GTGAAA_L001	1883	Salmonella_custom_v2

Salmonella enterica	281_GTGAAA_L001	2210	Salmonella_standard_v2
Salmonella enterica	281p2_CCGTCC_L001	2016	Salmonella_custom_v2
Salmonella enterica	281p2_CCGTCC_L001	2267	Salmonella_standard_v2
Salmonella enterica	285_GTGGCC_L001	2359	Salmonella_custom_v2
Salmonella enterica	285_GTGGCC_L001	2786	Salmonella_standard_v2
Salmonella enterica	285p2_GTGAAA_L002	2515	Salmonella_custom_v2
Salmonella enterica	285p2_GTGAAA_L002	2914	Salmonella_standard_v2
Salmonella enterica	286_GTTTCG_L002	1792	Salmonella_custom_v2
Salmonella enterica	286_GTTTCG_L002	2094	Salmonella_standard_v2
Salmonella enterica	286p2_GTCCGC_L002	1553	Salmonella_custom_v2
Salmonella enterica	286p2_GTCCGC_L002	1806	Salmonella_standard_v2
Salmonella enterica	287_CGTACG_L002	1953	Salmonella_custom_v2
Salmonella enterica	287_CGTACG_L002	2452	Salmonella_standard_v2
Salmonella enterica	287p2_ATCACG_L002	2120	Salmonella_custom_v2
Salmonella enterica	287p2_ATCACG_L002	2333	Salmonella_standard_v2
Salmonella enterica	289_GAGTGG_L003	2199	Salmonella_custom_v2

Salmonella enterica	289_GAGTGG_L003	2546	Salmonella_standard_v2
Salmonella enterica	289p2_CGATGT_L003	2061	Salmonella_custom_v2
Salmonella enterica	289p2_CGATGT_L003	2447	Salmonella_standard_v2
Salmonella enterica	298_ACTGAT_L003	2491	Salmonella_custom_v2
Salmonella enterica	298_ACTGAT_L003	2	Salmonella_standard_filter_v2
Salmonella enterica	298_ACTGAT_L003	3007	Salmonella_standard_v2
Salmonella enterica	298p2_TTAGGC_L003	1	Salmonella_custom_filter_v2
Salmonella enterica	298p2_TTAGGC_L003	1	Salmonella_custom_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	298p2_TTAGGC_L003	1	Salmonella_custom_filter_v2_ misclassified Salmonella
Salmonella enterica	298p2_TTAGGC_L003	1764	Salmonella_custom_v2
Salmonella enterica	298p2_TTAGGC_L003	1	Salmonella_standard_filter_v2
Salmonella enterica	298p2_TTAGGC_L003	1	Salmonella_standard_filter_v2_ Blast confirmed Salmonella
Salmonella enterica	298p2_TTAGGC_L003	2205	Salmonella_standard_v2
Salmonella enterica	349_TGACCA_L003	1449	Salmonella_custom_v2
Salmonella enterica	349_TGACCA_L003	1714	Salmonella_standard_v2
Salmonella enterica	349p2_CGATGT_L004	1948	Salmonella_custom_v2

Salmonella enterica	349p2_CGATGT_L004	2416	Salmonella_standard_v2
Salmonella enterica	37_CTTGTA_L001	1997	Salmonella_custom_v2
Salmonella enterica	37_CTTGTA_L001	1	Salmonella_standard_filter_v2
Salmonella enterica	37_CTTGTA_L001	2449	Salmonella_standard_v2
Salmonella enterica	372_CAGATC_L004	908	Salmonella_custom_v2
Salmonella enterica	372_CAGATC_L004	1072	Salmonella_standard_v2
Salmonella enterica	372p2_ACAGTG_L004	1682	Salmonella_custom_v2
Salmonella enterica	372p2_ACAGTG_L004	1989	Salmonella_standard_v2
Salmonella enterica	376_ACTTGA_L004	2179	Salmonella_custom_v2
Salmonella enterica	376_ACTTGA_L004	2509	Salmonella_standard_v2
Salmonella enterica	376p2_GCCAAT_L004	1916	Salmonella_custom_v2
Salmonella enterica	376p2_GCCAAT_L004	2238	Salmonella_standard_v2
Salmonella enterica	37p2_TAGCTT_L001	1755	Salmonella_custom_v2
Salmonella enterica	37p2_TAGCTT_L001	2242	Salmonella_standard_v2
Salmonella enterica	51_AGTCAA_L001	2965	Salmonella_custom_v2
Salmonella enterica	51_AGTCAA_L001	3465	Salmonella_standard_v2

Salmonella enterica	51p2_GGCTAC_L002	1712	Salmonella_custom_v2
Salmonella enterica	51p2_GGCTAC_L002	2193	Salmonella_standard_v2
Salmonella enterica	62_AGTTCC_L002	2405	Salmonella_custom_v2
Salmonella enterica	62_AGTTCC_L002	2869	Salmonella_standard_v2
Salmonella enterica	62p2_CTTGTA_L002	2107	Salmonella_custom_v2
Salmonella enterica	62p2_CTTGTA_L002	2301	Salmonella_standard_v2
Salmonella enterica	83_ATGTCA_L002	1946	Salmonella_custom_v2
Salmonella enterica	83_ATGTCA_L002	2304	Salmonella_standard_v2
Salmonella enterica	83p2_AGTCAA_L002	1981	Salmonella_custom_v2
Salmonella enterica	83p2_AGTCAA_L002	2224	Salmonella_standard_v2
Salmonella enterica	84_CCGTCC_L003	1971	Salmonella_custom_v2
Salmonella enterica	84_CCGTCC_L003	2227	Salmonella_standard_v2
Salmonella enterica	84p2_AGTTCC_L003	1346	Salmonella_custom_v2
Salmonella enterica	84p2_AGTTCC_L003	1558	Salmonella_standard_v2
Salmonella enterica	92_GTCCGC_L003	2289	Salmonella_custom_v2
Salmonella enterica	92_GTCCGC_L003	2610	Salmonella_standard_v2

Salmonella enterica	92p2_ATCACG_L003	2641	Salmonella_custom_v2
Salmonella enterica	92p2_ATCACG_L003	2947	Salmonella_standard_v2
Salmonella enterica	93_GTGAAA_L003	1986	Salmonella_custom_v2
Salmonella enterica	93_GTGAAA_L003	2338	Salmonella_standard_v2
Salmonella enterica	93p2_TTAGGC_L004	1674	Salmonella_custom_v2
Salmonella enterica	93p2_TTAGGC_L004	1930	Salmonella_standard_v2

Supplemental Table 4.1. Table of all a-priori metadata variables by type, used to represent AMD exposure for each sample. All metadata variables were included in the starting model for redundancy analysis and step-wise model selection was used to identify significantly associated variables with the resistome or microbiome.

Categorical	**Continuous: ADD**	**Continuous: in-feed ADD**	**Continuous: parenteral ADD**	**Continuous: Time**
Sampling time (Arrival, Rehandling)	Total ADD	In-feed ADD	Tetracycline ADD	Days-on-feed (DOF)
Feedlot ID	Total macrolide ADD	tetracycline in-feed ADD	Macrolide ADD	Days since first tx
Total ADD categories	Total tetracycline ADD	macrolide in-feed ADD	Betalactam ADD	Days since last tx
			Fluoroquinolone ADD	
			Phenicol ADD	
			Sulfonamide ADD	

Supplemental Table 4.2. Microbiome and resistome sequencing results for all 120 samples in the study.

Sample ID	Time	16S Raw paired reads	16S dada paired reads	16S dada filtered paired reads	AMR raw paired	AMR QC paired	AMR nonhost pair reads
IA1	Arrival	140704	49101	49101	0	0	0
IA10	Arrival	156473	39054	39054	0	0	0
IA11	Arrival	151460	38339	38339	0	0	0
IA12	Arrival	150609	38428	38428	0	0	0
IA13	Arrival	143444	41094	41094	6885106	6770930	4253208
IA14	Arrival	154019	46522	46516	0	0	0
IA15	Arrival	157856	41219	41195	13336698	12767024	11052170
IA16	Arrival	149815	35381	35381	18194281	17000763	15493555
IA17	Arrival	144616	36912	36912	0	0	0
IA19	Arrival	122561	33900	33887	22268461	21756652	13790848
IA2	Arrival	149420	43389	43389	3113837	3045174	2119555
IA23	Arrival	146311	36310	36310	23756211	23046236	16118824
IA24	Arrival	154089	44938	44938	12163267	11817805	8840755
IA27	Arrival	143781	35636	35636	0	0	0
IA29	Arrival	125618	47385	47376	17066935	16648607	12908731
IA3	Arrival	112200	32641	32641	0	0	0
IA30	Arrival	154982	39248	39248	9717604	9507432	6720600
IA32	Arrival	151680	38199	38199	10408872	10215177	6443623
IA33	Arrival	119859	37598	37598	10126581	9934145	6967346
IA34	Arrival	147866	51637	51637	13113540	12821912	9270086
IA35	Arrival	145085	44065	44065	19096271	18559078	15468273
IA38	Arrival	150252	25769	25769	16998160	16673804	10869383
IA39	Arrival	176334	45404	45404	16517780	16053959	12749966
IA4	Arrival	146792	38461	38461	0	0	0
IA40	Arrival	144712	38655	38655	17364153	16781053	14525989
IA41	Arrival	149849	35817	35817	18498535	17943350	14562290
IA43	Arrival	159747	38591	38591	16760676	16262775	13750391

IA44	Arrival	156667	36249	36249	0	0	0
IA45	Arrival	156332	38316	38316	16660388	16180543	13241837
IA46	Arrival	147495	33427	33427	5213762	5008062	3425972
IA47	Arrival	154382	31687	31687	11903784	11545274	8762648
IA49	Arrival	145691	32027	32027	13351222	12911886	9396998
IA5	Arrival	157802	55471	55471	18392226	17426039	14443544
IA50	Arrival	141149	38349	38349	0	0	0
IA51	Arrival	151384	32223	32223	9908348	9684955	6914998
IA52	Arrival	156177	30714	30714	17245535	16507286	13464331
IA53	Arrival	147787	34938	34938	8882200	8637459	5694920
IA54	Arrival	101543	24539	24539	13789787	13321816	10440449
IA55	Arrival	153294	33143	33143	13309205	12907268	8173028
IA56	Arrival	150321	32255	32255	17291458	16393869	14253414
IA57	Arrival	158787	40613	40594	8521160	8087587	6804680
IA58	Arrival	108365	25008	25008	19504110	18654819	16478960
IA6	Arrival	147032	40679	40679	0	0	0
IA60	Arrival	154506	34222	34222	19906054	18855981	16403164
IA61	Arrival	146339	28934	28934	14924071	14331765	12051725
IA62	Arrival	123446	20943	20943	21391921	20864639	15685088
IA63	Arrival	152160	31535	31535	12765482	12412086	9427508
IA64	Arrival	140853	43935	43935	20723988	19896626	15724592
IA65	Arrival	156439	42519	42519	14756702	14118605	11280303
IA66	Arrival	142067	44414	44414	14775648	14215617	11645321
IA67	Arrival	149940	44490	44490	12226238	11645032	9719303
IA68	Arrival	142477	54518	54518	13517390	13127797	11166604
IA7	Arrival	154469	37497	37497	0	0	0
IA70	Arrival	142826	47378	47378	18502435	18076053	13761314
IA71	Arrival	159977	30807	30786	15735681	15329113	11662085
IA72	Arrival	145745	42373	42373	17381438	16432568	13825225
IA73	Arrival	150886	47370	47370	14601711	13958795	11953633
IA74	Arrival	152937	39627	39614	17679730	16867383	14035003

IA75	Arrival	146274	49445	49445	15353917	14886637	12343325
IA76	Arrival	155443	50658	50658	14602926	14057362	10725495
IE1	Rehandling	156275	48140	48140	13806648	13089954	10591622
IE10	Rehandling	208020	41463	41463	21193832	20575827	16318227
IE11	Rehandling	146587	34103	34103	0	0	0
IE12	Rehandling	156433	36436	36436	0	0	0
IE13	Rehandling	106351	24930	24930	0	0	0
IE14	Rehandling	145371	50178	50178	0	0	0
IE15	Rehandling	146249	43976	43976	19685658	19086296	15729257
IE16	Rehandling	154160	53384	53384	13743238	13336915	10708659
IE17	Rehandling	171258	49272	49272	19059692	18440960	14689315
IE18	Rehandling	157239	46995	46995	16785654	16273548	12936134
IE2	Rehandling	159866	42133	42133	0	0	0
IE20	Rehandling	156935	47813	47813	18970820	18462433	13516461
IE24	Rehandling	159987	62936	62936	21504949	20868588	20858972
IE25	Rehandling	142662	42072	42072	17953607	17392219	17143446
IE28	Rehandling	143921	48452	48452	12257958	11903945	9458750
IE3	Rehandling	153190	48370	48370	0	0	0
IE30	Rehandling	159772	48234	48234	15142948	14667207	11884237
IE31	Rehandling	143922	36276	36276	17733210	17208323	13615863
IE33	Rehandling	150166	41449	41449	18042939	17531530	13210573
IE34	Rehandling	142918	53187	53187	17223598	16669061	14186473
IE35	Rehandling	140318	46463	46463	18471701	17908391	14793522
IE36	Rehandling	152489	38721	38721	13094397	12708668	10127699
IE37	Rehandling	150312	37912	37912	14245098	13786201	11435285
IE39	Rehandling	103412	33155	33155	19329174	18712524	15987733
IE4	Rehandling	199675	65270	65270	0	0	0
IE40	Rehandling	136214	35955	35955	14439790	13954691	11791550
IE41	Rehandling	142115	45533	45533	16977386	16555902	12188252
IE42	Rehandling	158718	49242	49242	16845632	16310340	14119822
IE43	Rehandling	105581	33674	33674	17868842	17259648	14449697

IE44	Rehandling	152088	40244	40244	15882170	15351204	12961009
IE45	Rehandling	145722	36864	36864	14104305	13658161	11154239
IE46	Rehandling	131935	43176	43176	9738094	9407972	8141444
IE47	Rehandling	156928	49793	49793	14187083	13752544	13632490
IE49	Rehandling	149310	42470	42461	0	0	0
IE5	Rehandling	111036	31287	31287	0	0	0
IE50	Rehandling	117846	31907	31907	0	0	0
IE51	Rehandling	140672	35001	35001	0	0	0
IE52	Rehandling	141109	53673	53673	17025545	16500001	13614725
IE53	Rehandling	151051	36111	36082	21767716	21195846	15718097
IE54	Rehandling	158845	47258	47253	17169236	16645905	13394455
IE55	Rehandling	148659	33338	33338	19797439	19232007	15068486
IE56	Rehandling	156905	34387	34387	24082109	23405466	18363671
IE57	Rehandling	157634	49062	49057	18138331	17585483	14895657
IE58	Rehandling	142634	42598	42598	8782167	8504175	6819840
IE59	Rehandling	154698	48224	48185	15375645	14820695	11710440
IE6	Rehandling	129039	38771	38743	0	0	0
IE62	Rehandling	155274	44864	44864	25239487	24531366	17114702
IE63	Rehandling	141720	48541	48541	14628392	14207255	11086453
IE64	Rehandling	143225	37533	37533	19488355	18840488	13938025
IE65	Rehandling	171377	61647	61647	16985106	15935624	13911750
IE66	Rehandling	148742	45584	45584	17682597	16739665	14219806
IE68	Rehandling	148005	46819	46553	17933154	16890376	14710387
IE69	Rehandling	146269	33207	32890	16221849	15522573	12186915
IE7	Rehandling	144425	39397	39274	0	0	0
IE70	Rehandling	121105	31187	31079	15702305	14948510	12072616
IE71	Rehandling	143137	43305	43305	15076613	14264924	11960739
IE73	Rehandling	148566	40010	39898	16343714	15362903	13085037
IE74	Rehandling	130808	39904	39876	20214701	19083866	16368529
IE75	Rehandling	158269	49488	49426	15658022	14880889	12403998
IE77	Rehandling	147345	38690	38584	17295179	16352131	13552227

Supplemental Table 5.1. Table of all a-priori metadata variables by type, used to represent AMD exposure for each sample. All metadata variables were included in the starting model for redundancy analysis and step-wise model selection was used to identify significantly associated variables with the resistome or microbiome.

Categorical	**Continuous: ADD**	**Continuous: in-feed ADD**	**Continuous: parenteral ADD**	**Continuous: Time**
Sampling time (Arrival, Re-handling)	Total ADD	In-feed ADD	Tetracycline ADD	Days-on-feed (DOF)
Feedlot ID	Total macrolide ADD	tetracycline in-feed ADD	Macrolide ADD	Days since first tx
Total ADD categories	Total tetracycline ADD	macrolide in-feed ADD	Betalactam ADD	Days since last tx
			Fluoroquinolone ADD	
			Phenicol ADD	
			Sulfonamide ADD	

Supplemental Table 5.2. Microbiome and resistome sequencing results for all 98 samples in the study.

ID	16S raw paired reads	16S dada paired reads	16S dada filtered paired reads	AMR raw paired reads	AMR QC paired reads	Time
PA1	205416	64523	64523	19245266	18651664	Arrival
PA10	200658	58042	58042	16313306	15760896	Arrival
PA12	202240	63877	63840	22845224	22021720	Arrival
PA13	211817	43185	43181	20358235	19667403	Arrival
PA14	203311	43525	43490	18127401	17746083	Arrival
PA15	203727	64735	64735	16724173	16198353	Arrival
PA16	201041	44562	44550	17983263	17478550	Arrival
PA17	203366	48937	48937	17306645	16676010	Arrival
PA18	203702	53561	53535	14501953	14287626	Arrival
PA19	94433	25997	25997	16832557	16319831	Arrival
PA2	209136	57305	57305	15899902	15353785	Arrival
PA20	217455	49004	48987	16220163	15688409	Arrival
PA21	195208	48435	48405	16806805	16293042	Arrival
PA23	207462	52298	52298	18113698	17597626	Arrival
PA24	124651	33303	33303	17358216	16856228	Arrival
PA25	214124	51944	51841	8739016	8480382	Arrival
PA26	205724	45251	45251	14393063	13950807	Arrival
PA27	143915	36997	36986	18623839	18043040	Arrival
PA28	219411	47830	47820	23948621	23160374	Arrival
PA29	205143	44340	44329	17465674	16820291	Arrival
PA30	217501	48811	48801	18184183	17532858	Arrival
PA31	219078	49514	49505	19729423	19133441	Arrival
PA32	210605	45963	45963	19838299	19220111	Arrival
PA33	206024	42699	42699	17747396	17181742	Arrival
PA35	202773	56452	56431	22449389	21653457	Arrival
PA36	205887	37837	37819	11782368	11362164	Arrival

PA37	205783	47474	47474	23922521	21946697	Arrival
PA38	201088	51123	51123	18012793	17301264	Arrival
PA39	201971	44042	44024	16101974	15339151	Arrival
PA4	145111	44552	44508	19966059	19279943	Arrival
PA40	190351	34962	34824	16032844	15434962	Arrival
PA41	204697	62960	62939	14235347	13735024	Arrival
PA42	204651	37223	37199	13237750	12830726	Arrival
PA43	204198	41821	41821	16540288	15978582	Arrival
PA45	211854	40277	40271	12191687	11719668	Arrival
PA46	115191	25441	25435	18517044	17969098	Arrival
PA47	202969	34288	34281	17649709	17129723	Arrival
PA48	212282	42536	42536	11188433	10835821	Arrival
PA49	214693	53228	53224	13878286	13461832	Arrival
PA50	148209	33657	33641	14745894	14301601	Arrival
PA7	208792	63902	63902	18534507	17872030	Arrival
PA8	200763	68276	68268	20241599	19392152	Arrival
PA9	212354	60190	60190	14877763	14364088	Arrival
PE10	166623	58715	58715	18605546	18051388	Rehandling
PE11	207772	66243	66243	17383250	16854961	Rehandling
PE12	211395	79485	79485	13215632	12766475	Rehandling
PE13	200454	59544	59544	14794685	14356244	Rehandling
PE14	211609	61091	61083	12709289	12337805	Rehandling
PE15	201707	50831	50828	14725154	14285630	Rehandling
PE16	202647	50785	50748	10966566	10570501	Rehandling
PE17	203761	43359	43341	11529527	11079686	Rehandling
PE18	204018	49260	49231	13209126	12765325	Rehandling
PE19	197577	55147	55141	18008957	17531127	Rehandling
PE2	215523	39327	39266	19594710	18833188	Rehandling
PE21	212627	47451	47451	13410196	13001729	Rehandling
PE22	219091	53854	53854	22258086	21526806	Rehandling
PE23	215521	55032	55032	21363130	20623950	Rehandling

PE25	103458	27701	27701	10303962	9996162	Rehandling
PE26	206272	47325	47325	11578677	11158964	Rehandling
PE27	183469	46817	46814	13006223	12504874	Rehandling
PE28	210838	60535	60523	13680789	13192312	Rehandling
PE29	95639	26442	26329	12550326	12161506	Rehandling
PE3	204476	33438	33369	20392844	19805318	Rehandling
PE30	206102	54367	54338	18434389	17738591	Rehandling
PE31	212619	50124	50091	22441037	21761975	Rehandling
PE32	140657	39556	39556	25456702	24669873	Rehandling
PE33	211518	51242	51242	20318226	19625304	Rehandling
PE34	206879	63179	63179	14988063	14453478	Rehandling
PE35	217120	59860	59860	15013056	14446075	Rehandling
PE36	179664	53362	53347	11422713	11043192	Rehandling
PE37	200111	57352	57352	14473202	13980445	Rehandling
PE38	203374	56975	56905	16334138	15511640	Rehandling
PE39	206855	81136	81128	16984631	15928258	Rehandling
PE4	200890	38365	38365	16603644	16097593	Rehandling
PE41	217312	47322	47201	10757992	10399014	Rehandling
PE42	201411	49164	49164	13160352	12264992	Rehandling
PE43	213548	61154	61154	17373341	16315254	Rehandling
PE44	216177	47389	47369	19137749	18472635	Rehandling
PE45	214722	58763	58612	18328123	17777958	Rehandling
PE46	214006	63965	63942	21301375	20599212	Rehandling
PE47	209128	47782	47730	14405735	13953678	Rehandling
PE48	210815	52862	52837	21461477	20781171	Rehandling
PE5	201779	35415	35415	14315547	13893309	Rehandling
PE50	213653	65150	65150	15033985	14271825	Arrival
PE51	175362	44874	44692	19237624	18579081	Rehandling
PE52	204758	44950	44919	19299089	18581227	Rehandling
PE7	211284	72073	72073	17873184	17314678	Rehandling
PE9	204642	52234	52226	14920293	14468008	Rehandling

PS1	205425	54226	54209	20613785	19903805	Shipment
PS11	188861	57893	57881	15709036	15163390	Shipment
PS12	203682	51993	51964	15199411	14473307	Shipment
PS2	206228	56843	56835	13148475	12683065	Shipment
PS3	129087	36627	36614	13321580	12814782	Shipment
PS4	205454	54612	54579	14531093	13878706	Shipment
PS5	219918	46496	46414	18304021	17589335	Shipment
PS6	218821	58207	58207	15993215	15456393	Shipment
PS8	213157	71908	71908	11279681	10614287	Shipment
PS9	214865	49111	49043	6192389	5827204	Shipment

Supplemental Table 6.1. Metadata variables for all 16 individual retail ground beef samples in the study.

ID	Label	Packaging samples	Lean	Blinded Store
FC_112	RWA	Vacuum(N=16)	90	Store6
FC_113	RWA	Vacuum(N=16)	90	Store2
FC_122	RWA	Vacuum(N=16)	85	Store4
FC_123	RWA	Vacuum(N=16)	90	Store4
FC_133	RWA	Vacuum(N=16)	92	Store1
FC_134	RWA	Vacuum(N=16)	80	Store1
FC_142	RWA	Vacuum(N=16)	93	Store3
FC_143	RWA	Vacuum(N=16)	85	Store3
FC_161	CONV	Chub(N=8)	80	Store1
FC_162	CONV	Chub(N=8)	93	Store1
FC_172	CONV	TrayOverwrap(N=6)	80	Store3
FC_177	CONV	TrayOverwrap(N=6)	93	Store5
FC_181	CONV	StoreGrind(N=2)	90	Store1
FC_186	CONV	Chub(N=8)	80	Store5
FC_195	CONV	Chub(N=8)	85	Store5
FC_199	CONV	TrayOverwrap(N=6)	93	Store3

Supplemental Table 6.2. Resistome sequencing results for 16 individual retail ground beef samples and their corresponding duplicate sample (N=32) which were processed with a diluted AMR bait-enrichment protocol.

ID	Label	Dilution	Blinded Store	Raw paired reads	QC filtered reads	nonhost filtered reads	deduped SNP confirmed counts
FC_112	RWA	None	Store6	38,190,379	36,762,458	28,728	100
FC_112	RWA	Half	Store6	8,544,874	8,249,150	5,538	673
FC_113	RWA	None	Store2	34,644,683	33,386,392	167,830	9,628
FC_113	RWA	Half	Store2	49,990,544	48,466,450	172,974	3,777
FC_122	RWA	None	Store4	42,523,674	41,038,940	98,088	19,089
FC_122	RWA	Half	Store4	59,222,209	57,116,177	146,919	12,514
FC_123	RWA	None	Store4	37,062,617	35,657,217	20,149	80
FC_123	RWA	Half	Store4	19,036,211	18,357,780	8,595	271
FC_133	RWA	None	Store1	41,617,299	39,968,343	65,852	628
FC_133	RWA	Half	Store1	15,200,330	14,620,247	27,615	6,833
FC_134	RWA	None	Store1	46,514,396	44,608,097	37,377	984
FC_134	RWA	Half	Store1	34,804,270	33,470,715	52,034	1,161
FC_142	RWA	None	Store3	34,983,723	33,538,488	78,496	5,303
FC_142	RWA	Half	Store3	41,672,381	40,236,202	81,513	1,342
FC_143	RWA	None	Store3	45,442,442	43,529,554	51,204	627
FC_143	RWA	Half	Store3	41,210,567	39,588,437	40,161	448
FC_161	CONV	None	Store1	43,515,676	41,896,445	728,548	51,868
FC_161	CONV	Half	Store1	67,058,362	64,579,770	824,877	34,143
FC_162	CONV	None	Store1	54,399,943	52,389,506	165,437	2,792

FC_162	CONV	Half	Store1	50,459,468	48,697,965	126,973	1,619
FC_172	CONV	None	Store3	45,710,363	44,090,217	1,410,649	2,147
FC_172	CONV	Half	Store3	39,790,090	38,407,079	1,104,944	20,120
FC_177	CONV	None	Store5	40,915,803	39,211,802	41,483	10,290
FC_177	CONV	Half	Store5	54,833,376	52,650,008	45,942	5,686
FC_181	CONV	None	Store1	50,492,224	48,832,439	66,376	266
FC_181	CONV	Half	Store1	44,527,155	43,105,651	50,912	135
FC_186	CONV	None	Store5	46,960,256	45,204,364	736,031	30,259
FC_186	CONV	Half	Store5	54,816,543	52,698,955	1,069,355	25,356
FC_195	CONV	None	Store5	48,943,119	47,226,073	197,321	4,722
FC_195	CONV	Half	Store5	43,918,073	42,408,126	175,310	10,623
FC_199	CONV	None	Store3	33,252,253	31,859,552	144,585	3,088
FC_199	CONV	Half	Store3	52,680,331	50,601,018	184,951	1,350

Supplemental Table 6.3. Microbiome sequencing results for 16 individual retail ground beef samples.

ID	Label	Blinded Store	Raw paired reads	QC filtered reads	taxa classified reads	filtered taxa reads
FC_112	RWA	Store6	206239	191802	46543	46355
FC_113	RWA	Store2	202019	188456	45206	45039
FC_122	RWA	Store4	213646	197900	40696	40451
FC_123	RWA	Store4	210004	184008	75558	75204
FC_133	RWA	Store1	212618	198432	30161	30078
FC_134	RWA	Store1	219822	198879	19258	18978
FC_142	RWA	Store3	187892	174760	33405	33298
FC_143	RWA	Store3	155859	143060	13081	12877
FC_161	CONV	Store1	180094	167597	42412	42122
FC_162	CONV	Store1	200365	188043	36662	36519
FC_172	CONV	Store3	100939	94569	16556	16363
FC_177	CONV	Store5	178300	162800	30436	28153
FC_181	CONV	Store1	216235	203920	38858	38709
FC_186	CONV	Store5	205446	192819	42174	41987
FC_195	CONV	Store5	207711	192730	46749	46658
FC_199	CONV	Store3	213344	203177	32762	32708