DISSERTATION

SHOTGUN METAGENOMIC AND 16S rRNA GENE SEQUENCING INVESTIGATION OF ANTIMICROBIAL RESISTANCE IN THE BEEF SUPPLY CHAIN

Submitted by

Margaret Weinroth

Department of Animal Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, CO

Spring 2019

Doctoral Committee:

Advisor: Keith E. Belk

Jessica L. Metcalf
Paul S. Morley
Hua Yang

ABSTRACT

Antibiotic resistant bacterial infections are a growing public health issue. In food producing animals, there is a concern that using antibiotics will increase the risk of antibiotic resistance (AMR) on meat products, thus increase consumers' risk of acquiring AMR infections via meat consumption. However, in order to mitigate these risks, there are several areas of livestock production where more data are needed to understand current communities of AMR genes (resistomes) and how antibiotics influence associated cattle microbiomes. In order to obtain more information about these communities, three experiments were conducted: (A) a characterization of market cow resistome and microbiomes, (B) the effect of region and tylosin exposure on feedlot cattle, and (C) the characterization of liver purulent material from feedlot cattle. All three of the studies were conducted utilizing shotgun metagenomics for resistome analysis and 16S rRNA amplicon sequencing for microbiome analysis. Both culture-free methods used in these studies were chosen because AMR is an ecological concern in microbial communities and resistance genes are not exclusively harbored in culturable bacteria. In the first study, market cows were selected as the population of interest because, while a fifth of beef produced in the U.S. is from market cows, past studies have been more focused on feedlot cattle. Findings from the market cows study highlighted that the resistome of bovine trimmings was impacted by slaughter facility over the production system the cows were raised in. In the second study, cattle were raised in three different geographical regions, and within each geographical region pens were either fed or not fed tylosin—a common macrolide antibiotic for the prevention of liver abscesses. In addition to the scoring of liver abscesses in these pens of cattle at harvest, pen floor feces and soil were collected. These data revealed no significant differences in

resistome composition between different antibiotic group exposures, but geographical region affected the resistome. Finally, aliquots of liver purulent material of a subset of cattle in the aforementioned experiment were collected for characterization in the third study. While tylosin is used to reduce liver abscesses in feedlot cattle, and it has been postulated that *fusobacterium necrophorum* is the causative organism of liver abscesses, the mode of action in which tylosin reduced abscesses is not known, nor has *fusobacterium necrophorum* introduction into livers resulted in 100% incidence of liver abscesses. To gain a more complete understanding of the cause of liver abscesses, purulent material was evaluated with 16S rRNA amplicon sequencing. *Fusobacterium* was identified in all abscesses along with many other phyla, demonstrating a polymicrobial bacterial community. All three of these studies further contributed to the understanding of resistome and microbiome dynamics as a result of antibiotic exposure. Across studies, geographical region and facility of slaughter were seen to contribute more to resistome composition changes in both feces and colon content and bovine trimmings than antibiotic exposure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Metagenomics of Food Commodities: Meats and Poultry[1]

Meat and poultry products, the animals that produce them, and the environments they are processed in are biologically diverse and contain many (thousands) species of bacteria, viruses, and fungi arranged in complex living environments. These organisms work together and share resources (such as metabolites), much like a community. Within this community, each organism has its own genome, yet some genes are ubiquitous across the community, while others are more specialized, present in only a few or one single species in the community. Together, the genomes and genes within this community are known as the pan-genome (Soucy et al., 2015). By looking at samples through the context of community-wide interactions, instead of a few select organisms, we can obtain a more holistic model of what is occurring in a given microbial community. This ecological way of thinking can be incorporated into any biological research, including meat and poultry production.

However, even with advancements in microbiological culture techniques, many organisms within and around meat and livestock products will likely never be cultured in a laboratory. This issue is known as the "plate count anomaly", which leads to estimates that from 76% (Suau et al., 1999) to upward of 99.9% (Youssef et al., 2015) of cells within samples cannot be found with culture plating or enumeration procedures alone. While the vast majority of cells will never be cultured in a laboratory setting, these unculturable bacteria still interact within biological communities, contribute to meat safety and quality, and affect the dissemination of genes of concern to public health, such as virulence factors. To overcome culture-dependent

[1] Authors: Margaret D. Weinroth, Noelle R. Noyes, Paul S. Morley, Keith E. Belk; this work will also be published in 5th edition of *Food Microbiology: Fundamentals and Frontiers*

methods, researchers have started integrating culture-free approaches into traditional studies with techniques such as metagenomics, transcriptomics, proteomics, metabolomics, and whole genome sequencing.

This chapter will specifically address the science of metagenomics, and how it is used in meat and poultry production to advance food safety and quality.

Overview of Metagenomics

Metagenomics is defined as the direct analysis of genomes contained within an environment (such as a processing facility, animal model, clinical samples, etc.) (Thomas et al., 2012); such studies can be sequence- or function-based. The term metagenomics was first coined in 1998 by Handelsman et al. (Handelsman et al., 1998) in reference to the functional analysis of all of the genomes within a soil sample. Sequence-based studies deal with the direct genetic analysis of genomes contained within environmental samples without a culture step (Thomas et al., 2012). Function-based studies, on the other hand, are conducted by cloning large portions of screened DNA into a host (Culligan and Sleator, 2016). Here, we focus on sequence-based metagenomics.

History and Process

As a precursor to modern sequence-based DNA sequencing, originally, whole community DNA analysis was conducted by cloning environmental DNA into recipient cells (Simon and Daniel, 2011). This method was eventually replaced by methods that directly isolate nucleic acids from the environment and sequence them; the first such method was Sanger Sequencing in 1977 (Sanger et al., 1977). Then, in the mid 2000's, next-generation sequencing (NGS) was introduced with the adoption of pyrosequencing technologies. Next Generation Sequencing (NGS) uses the concept of sequencing-by-synthesis, which allows for parallel sequencing

reactions to occur simultaneously, increasing sequencing ability from 96 reactions at a time to several hundred thousand (Schuster, 2008). The introduction of NGS brought sequencing costs down by 10,000 fold and greatly reduced labor requirements, which were intensive for Sanger-based sequencing. This technology was first commercialized by 454 Life Science and improved upon by a method known as bridge sequencing—first marketed by Solexa, which was then acquired by Illumina (Heather and Chain, 2016). As a result, DNA sequencing became a tool that could be implemented in many laboratories to address more applied topics, including those associated with agricultural production and processing (Muir et al., 2016). While NGS DNA sequencing was initially used for single genomes, it eventually became feasible to sequence entire communities of microbial DNA without any need for culture- or PCR-based isolation.

Sequencing a non-enriched sample from an environment of interest (whether that be a processed meat product, feces from a food producing animal, a production facility that packages ready-to-eat products, human clinical samples from a foodborne outbreak, or many other environments of interest to a food microbiologist) provides insights into the structure, composition and function of the microbial community – which, together, can be referred to as the microbial ecology of the sample. The microbiome is defined as the entire population, and the makeup of the community of organisms in a given sample or animal. In order to investigate microbial ecology (the interactions of the microbiome with the environment and host), many steps need to be performed – from study design to molecular biology to bioinformatics to statistical analysis.

Before sampling is conducted, an experimental design is constructed to meet research objectives. While many metagenomic studies currently conducted in meat and poultry science are more explanatory in nature (i.e., quantification of what organisms or genes are present in an

environment and in what quantities), a shift to hypothesis-driven research is beginning to occur. After the experiment is structured, sampling occurs as aseptically as possible, with special attention paid to not introducing foreign DNA contamination into the sample. From there, DNA is extracted and prepared for sequencing through PCR amplification, 'cleanup' steps, and ligation of adapters onto the DNA so that the sequencers can recognize the reads (known as library preparation). Fragmentation, breaking DNA into smaller pieces, also is conducted during library preparation, and is a key differentiator between short-read sequencing where this occurs and long-read sequencer where it does not. The manner in which samples are handled (including the specific kits and protocols used to extract DNA and conduct library preparation) can both introduce bias into sequencing and downstream analysis and must be considered carefully in the study design and planning stages (van Dijk et al., 2014). A major breakthrough in the preparation of libraries for sequencing was the ability to multiplex samples through use of barcoding. Barcoding involves the addition of a known, sample-specific nucleotide sequence onto each fragment of DNA, followed by pooling of DNA from multiple samples. This technique greatly reduces sequencing costs by enabling simultaneous sequencing of multiple samples, while maintaining the ability to match DNA sequences back to their originating sample during post-sequencing analysis.

It is important to note that one of the main differences between a whole-genome and a metagenomic pipeline is the absence of an enrichment step in the metagenomic pipeline. While single-genome workflows enrich for specific bacteria (through selective culturing and broths), a main objective of metagenomics is to look at a non-enriched view of the entire community. Thus, no steps are taken to differentially increase abundance of specific species, which would not allow for an accurate relative abundance estimations. Sample-processing steps must be as

unbiased as possible so as not to artificially perturb the composition of the microbial community (van Dijk et al., 2014).

Once DNA has been extracted, fragmented and a library is generated, sequencing of DNA is either achieved through short-read or long read sequencing. To-date, metagenomic studies have mostly utilized short-read sequencing due to lower costs, greater general accuracy and higher output compared to long-read sequencing technologies. The primary disadvantage of short-read technology is the inability to assign what genes or features of interest belong to what organism due to the shearing of DNA for library preparation.

The process of sequencing DNA produces fastq files that contain the DNA sequence of each fragment of DNA, as well as the sequencing quality score for each nucleotide. These short fragments are referred to as "reads", and each read within fastq files have unique identifiers. With these raw sequence reads as a starting point, bioinformatic analyses can begin. The first step of the bioinformatics pipeline is typically quality control of the sequence data, which involves removal of library adapter sequences; removal of low-quality nucleotides on the ends of reads; and removal of reads that have overall low quality or are too short. For analyses in which bacterial community is of main interest (e.g., a study on the shelf-life of meat), reads that belong to the host (e.g., *Bos taurus* for beef or *Gallus gallus* to reflect chicken DNA) can be removed as part of the quality control process.

After quality control is complete, assembly and/or alignment of sequence reads can begin. Assembly involves stringing the short reads back together into longer sequences called "contigs", typically using de Bruijn graph-based or overlap-layout-consensus algorithms (Miller et al., 2010). Due to the nature of metagenomic DNA, assemblies tend to be highly "fragmented"; i.e., with numerous short contigs, especially compared to assembly of single-

genome DNA sequence data. Once contigs are assembled, they typically are compared to existing databases using matching algorithms such as BLAST (Altschul et al., 1990) in order to identify sequences of interest.

Alignment, on the other hand, allows bypassing of assembly and attempts to match unassembled sequence reads directly to existing databases, typically using algorithms such as the Burrows-Wheeler-Aligner (Li and Durbin, 2010). Whether using assembled or unassembled data, the process of matching DNA sequences to existing databases allows for identification and quantification of genes and/or organisms of interest within the metagenomic data (note that, in the case of 16S sequencing which is discussed below, only phylogenetic classification can occur while shotgun metagenomic sequencing provides a much broader scope of genomes present in the sample, along with functional and gene differentiation).

Once genes and/or organisms of interest have been identified and counted, descriptive and formal inferential statistical tests can be conducted in conjunction with previously collected metadata (i.e., descriptive data associated with the samples such as location, sample matrix, environmental characteristics, etc.). Specific comparisons made during statistical analysis depend largely on study design, although basic descriptions such as numbers of reads, read quality, and level of host contamination are typically reported for all metagenomic studies. After formal comparisons have been conducted, raw data and critical output files must be stored on a secure server or backed up in another manner, such as in the "cloud". As with other sequencing studies, metagenomic datasets are often deposited as raw sequence data onto public repositories such as the National Center for Biotechnology Information Sequence Read Archive (NCBI SRA) (Leinonen et al., 2011), as this is frequently a requirement for publication in high-quality scientific journals.

Terminology used in discussions of metagenomic studies currently is in flux, and oftentimes 16S microbiome studies and/or extensive sequencing of a single species are referred to as “metagenomics”. It is important to note that, from a purist point of view, the term metagenomic means “genome of genomes” and thus refers to whole-community, unrestricted sequencing of all the genomes in a sample (known as shotgun metagenomics). Alternatively, 16S studies involve PCR amplification and sequencing of the 16S rRNA gene as a marker to investigate the taxonomic bacterial composition of a sample (Jovel et al., 2016). Because shotgun is unspecific in what piece of the bacterial genome is amplified and sequenced, it can be used to look at microbial community composition, as well as other features of interest, such as specific genes and detailed metabolic and functional profiles (Parks and Beiko, 2013). On the other hand, 16S uses primers that anneal to conserved genomic regions that contain segments of hyper-variability (known as the V1 to V9 variable regions). The 16S gene itself is specific to bacteria and archaea and is highly conserved between species, making it a useful target to capture most of the bacteria within a sample; the hyper-variable regions within the 16S gene, on the other hand, tend to differ between bacterial genera and/or species, and thus can be used to reconstruct the taxonomy of the bacteria and archaea within the sample.

While much less expensive than shotgun metagenomics, there are some drawbacks specific to 16S analysis that are caused by amplifying only a portion of the genome (as opposed to unrestricted sampling of the entire genome, as in shotgun studies). Namely, 16S reads cannot be assigned to taxa at high resolution, such as at the species or stain level (Jovel et al., 2016). Additionally, because only the 16S gene is targeted and sequenced, use of 16S sequencing does not allow direct functional analysis of the communities being studied. Some pipelines exist that attempt to perform correlative analysis between the taxonomic composition of a sample and its

potential functional capacity (e.g., potential metabolic pathways). Some of these pipelines use curated databases of both genomes and so-called "marker genes", the latter of which have been validated to correlate to both specific bacterial taxa and specific bacterial functions. It is important to note that the accuracy of such correlative analyses is heavily debated in the scientific literature (Segata et al., 2012).

Bioinformatics

A unique component of analyzing large biological datasets is reducing gigabytes (or more) of data down into manageable, useful information. Bioinformatics is defined as the application of computational tools to the capture and interpretation of biological data (Bayat, 2002). Bioinformatics is multidisciplinary in nature as it combines expertise of fields such as computer science, statistics, and biology to understand and interpret the information contained in large datasets.

In shotgun metagenomics, there are many tasks that must be automated through shell scripts and algorithms in order to make them efficient. These tasks can range from simple to more complex, but all of the tasks must be well understood in order to verify that the bioinformatics tools are being used appropriately. An example of a simple, but essential, tool is read trimming and filtering (i.e., quality control or QC, see above). After raw reads are produced by the sequencer, removal of sequencing adaptors and low quality nucleotide reads is an essential task. While this would be easy to do manually on one read, a typical study has billions to trillions of reads to process. Tools like Trimmomatic (Bolger et al., 2014a) provide free, open source scripts for public use. Published tools can be focused on a specific task, such as trimming or alignment, or can be in the form of "pipelines" that wrap multiple tools into one bioinformatic workflow. Wrapping multiple tools into one computational workflow not only helps with speed,

but also with data reproducibility. Galaxy (Afgan et al., 2016) is an example of a Reproducible Research System (RRS), which is an environment for executing and tracking computational analyses (Goecks et al., 2010). The advantage of these types of systems is that they not only keep the analysis organized, but can also be published alongside research findings for reproducibility.

Special consideration: on target reads. When preparing to conduct an analysis, evaluation of whether or not there are enough sequences in relation to the sample type, and enough genetic read diversity, to conduct a statistical analysis. One way to assess this is to perform rarefaction of the sequence data to produce a rarefaction curve. A rarefaction curve traditionally displays sampling depth on the x-axis and features (species, genes, etc.) or a measure of richness on the y-axis. If sampling depth (i.e., the number of sequences attributed to each sample) is appropriate, the curve constructed for the sample (or group of samples) should level out when an appropriate sampling depth is reached. The depth of sequencing that needs to occur differs by bacterial community type as a function of diversity (i.e., the more diverse the matrix, the more reads are needed). While deep sequencing (i.e., sequencing a genome many times, sometimes hundreds or even thousands of times, and generation of a high number of unique reads) would be the most ideal for research, the cost makes ‘unlimited’ sequencing impossible. As a result, pre-study work with a smaller subset of the sample type of interest, or relaying on previously published work, needs to be considered before a sequencing depth is targeted.

While appropriate sequencing depth is a concern, the number of reads that map to a feature of interest (known as a ‘hit’) must also be considered. For example, in a study conducted on beef cattle production (Noyes et al., 2016) that sampled both feces and meat, only 1.5% of total raw reads from feces aligned to the bovine host genome and were removed from the sequence data, while 99.6% of sequences in the meat were associated with the bovine genome.

Their study was conducted to evaluate AMR genes in the samples, yet across all samples, only 0.04% of all raw sequences were attributed to one of such AMR genes. As a result, 99.6% of the sequences in this study did not contribute to the research objectives and could be considered off-target. This is not an uncommon occurrence in metagenomic studies in which the objective is to investigate a subset of genes within a wider population. There has been some work to combat this issue, including use of both functional metagenomics (not discussed in this chapter) and targeted sequencing. Targeted sequencing involves specifying features of interest and designing 'baits' for these targets (Sulonen et al., 2011). Once baits have been designed, they can be used to enrich for specific sequences within the metagenomic DNA via magnetic streptavidin beads, resulting in more on-target hits than a sample for which targeted sequencing is not used.

Special consideration: dealing with different library sizes. There are many decisions that must be made to minimize bias when conducting a metagenomic study. These decisions start with experimental design, and carry through to data presentation. One decision that must be made is how to "normalize" sequence data. Weiss et al. described normalizing as, "the process of transforming the data in order to enable accurate comparison of statistics from different measurements by eliminating artifactual biases in the original measurements (Weiss et al., 2017)." While library preparation and sequencing technologies have improved dramatically in short periods of time, there still can be an issue with uneven sequencing of DNA from samples. This results in sequence read datasets of differing sizes. For example, equal volumes of DNA extracted from two samples can be sequenced on the same sequencer at the same time, but one may generate 20 million reads while the other may generate 40 million reads.

While differences in the number of sequences obtained from different samples are common, they are rarely a reflection of true biological differences, but more of sequencing

inconsistency. A few factors that affect read number include raw DNA allocation when pooling DNA for sequencing, sequencing protocol and sample matrix, and sequencer efficiency—such as variation in output from lane to lane on the same sequencer (Aird et al., 2011) . Because it is currently impossible to ensure completely uniform sequencing depth, post-sequencing tactics must be used to normalize for the uneven numbers of reads generated per sample.

Two common methods of normalizing are rarefying and scaling; both provide advantages and disadvantages. Rarefying is the process of randomly sampling without replacement from each sample up to a common count (Weiss et al., 2017). For example, if one sample had 1,000 reads and another had 2,000 reads, and the rarefying depth was set at 1,000 reads, the smaller of the two samples would be completely represented because all 1,000 reads would be sampled. From the 2,000 read sample, 1,000 reads would randomly be drawn without replacement and the rest discarded. This would result in an equal number of sequences for each of the two samples (i.e., 1,000), and these reads would be carried forward into statistical analyses. The minimum number for rarefying could be selected as the lowest number of reads per sample within the dataset; alternatively, a higher number could be chosen, resulting in exclusion of samples with a lower number of reads. Scaling, on the other hand, is multiplication of count by a number or proportion. While total sum scaling (TSS) divides counts by total number of reads, a more sophisticated technique known as cumulative-sum scaling (CSS) builds on this method such that counts are divided by the cumulative sum of counts up to a percentile determined by the data (Paulson et al., 2013a). Another approach to dealing with count data is a logarithmic-ratio transformation. However, because a log transformation can only be used with positive numbers, pseudocounts must be added to zeros (i.e., usually adding a 1 to each observation), an issue that can result in bias (Martín-Fernández et al., 2015).

Rarefying effectively mitigates varied sample read count size and is especially versatile in handing small and uneven library sizes (Weiss et al., 2017). Disadvantages to rarefying include "throwing away data" (i.e., using only a small proportion of total reads) and adding an additional uncertainty to the data (McMurdie and Holmes, 2014). In contrast, CSS and other scaling methods utilize all counts in the sequencing data by adjusting them by a proportion. However, scaling also has problems associated with it. It can over- or under-estimate rare features in samples (Weiss et al., 2017). These issues can be illustrated by looking at two library sizes, one of 10,000 counts and another of 100,000 counts. In the small library, overestimation of abundance can occur (11). For example, if a feature were present in both samples at a minimal number, the scaling factor applied to the smaller library may overestimate a feature that is, in fact, very rare in the population.

Normalizing currently is a necessary tool due to vastly different sequencing library sizes among samples. Unfortunately, there is no perfect method that allows all data to be evaluated while avoiding over- or under-estimations of features of interest. However, there are tools being developed that incorporate more robust models to handle these types of variably-sized data (Chouvarine et al., 2016). This is an area that has seen recent growth and will continue to evolve as bioinformaticians and statisticians develop models that more correctly handle count data.

Data Presentation

Metagenomic data are multivariate, and studies routinely generate gigabytes and even terabytes of data. While bioinformatic tools reduce these large datasets to more manageable sizes, the way in which metagenomic results are reported and visualized often differ when compared to traditional culture-based reporting methods; depending on objective(s), there are many visualizations that can be used. A taxonomy plot, or 100% stacked bar graphs, can

represent individual sample diversity, or the taxonomy of groups of samples that share common phenotypic characteristics. Ordination plots (including Principle Component Analysis [PCA] and Non-metric multidimensional scaling [NMDS] plots) are used to visualize calculated orthogonal distances between samples based on some mathematical associations in multidimensional space that account for sequentially-declining amounts of variance, which typically reflect differences in the composition of the microbiome. This allows for visualization of "like" and "unlike" samples, i.e., clustering. Heatmaps can be used to illustrate feature abundances (such as a species of interest or specific genes or classes of genes) across samples or groups of samples, again with the goal of visualizing clusters of like samples.

Phylogenetic relationships are 'trees' built on how similar or dissimilar species or communities are. Uses for these trees can include illustrating the full phylogenetic diversity of a sample, or comparing a subsection of data, such as the diversity of species within a genus. Network analysis is the visualization of associations between genes or other features of interest. Co-occurrence of these features is something that can be demonstrated with a network analysis, as well as associations. While these methods of visualization are a starting point for data interpretation, there are numerous, more robust hypothesis-testing methods such as ZIG models (Paulson et al., 2013c), PERMANOVA (Tang et al., 2016), and Bayesian Network Analysis (Hobbs et al., 2016) that are used in metagenomics, but that are outside the scope of this chapter.

Advantages of Metagenomics

In addition to the relative reduction in cost in the last decade, and the ability to study uncultured organisms, metagenomic methods have other advantages as well. These include the ability created to assess community ecology and to study unknown pathogens of interest. The most touted advantage of metagenomics is the ability to look at an entire community of bacteria.

Culture-based techniques often use indicator organisms, such as generic *E. coli* or aerobic plate counts, as markers for pathogenic bacteria. While indicator organisms are used as surrogates for indicating potential of pathogen presence as a fraction of a larger population, some studies have shown only limited direct relationships between indicators and pathogens (Harwood et al., 2005). Instead, in some cases, use of indicator organisms rests on the assumption that, if a reduction is induced or occurs in the larger microbial population, the pathogenic sub-population also will decline in magnitude accordingly (Brown et al., 2000). Metagenomics, on the other hand, can be used to directly assess presence and abundance of organisms in a microbial population, as well as presence of specific genes of interest. While metagenomics has primarily been used to develop inferences concerning bacteria, the technique also allows for in-depth investigation of the virome (all the viruses in a given environment), the mycobiome (all the fungi in an environment), and the plasmidome (all the plasmids). This is of special interest in food microbiology, where processes like shelf life and aging also are heavily influenced by fungal growth.

In addition to identifying and quantifying organisms within a given microbial community, interactions between organisms can be studied. To assess relationships between features of interest (whether genes, genomes or sets of genomes) in a bacterial population, co-occurrence and co-exclusion can be studied in the context of network analysis. Co-occurrence and co- exclusion relationships look at whether a specific feature of interest occurs at a higher rate alongside another feature; for example, two genes that are on the same plasmid would have a higher co-occurrence than two chromosomal genes from different species of bacteria. Network analysis can provide insight into inter-taxa associations to understand symbioses between community members; this allows more complex pattern discovery than traditional diversity

metrics (Barberán et al., 2012). Understanding associations between different microorganisms within a bacterial community can guide researchers towards specific targets for alteration or manipulation of the community—a systems approach to food.

Disadvantages of Metagenomics

While NGS has lowered sequencing costs, there are still financial barriers. These barriers are not just limited to generation of sequence data, but also to computational capabilities, data storage and, in some cases, the need to hire a bioinformatician and/or laboratory personnel familiar with NGS. As sequencing costs drop rapidly, costs associated with data analysis and storage are forecast to occupy a larger proportion of total experimental costs. After overcoming financial costs associated with the technique, there are hurdles related to metagenomics being a relatively new method for ecological investigation; namely shortcomings associated with public databases, and availability and usability of analytical tools.

With respect to public databases used in metagenomic analysis, there are several difficulties that researchers may encounter, including 1) incomplete databases, 2) mis-annotation, and 3) inconsistent nomenclature. While the rise in NGS has helped fuel expansion of many databases, there are still features and organisms that have yet to be entered into a database. To combat this issue, one method of de novo discovery of novel genes or organisms is called *de novo* assembly (Li et al., 2010).The issue of un-annotated sequence data is especially prevalent for fungi; while public databases exist, they are generally a few years behind those established for bacteria (Mayo et al., 2014). This is especially challenging to food researchers studying shelf life and fermentation, where fungi play an important role in bacterial communities.

The second issue surrounding databases is the possibility of mis-annotation (i.e., that a gene or species of interest is erroneously attributed to an incorrect species due to an error in the

database). When mis-annotation occurs, the only remedy is careful curation of the database, which is time and labor intensive.

Finally, there is the concern over inconsistent nomenclature/taxonomic grouping "rules." This can be addressed by instituting standards for the naming of features or organisms. An example of this is found in antibiotic resistance gene naming where a gene needs at least a 2% difference in nucleotide makeup than its closest relative for it to be designated as a 'new' gene (Hall and Schwarz, 2016).

Metagenomic studies require merging of computational and biological knowledge. Historically, these disciplines (i.e., computer science and the life sciences) have been largely separate. As life science researchers increasingly utilize bioinformatic software, the user-friendliness of these tools continues to improve (Mayo et al., 2014). Use of 16S sequencing, which has a longer track record than more costly shotgun metagenomic sequencing, can be analyzed using two widely-available platforms (MOTHUR (Schloss et al., 2009) or QIIME (Caporaso et al., 2010a)), both of which have graphical user interfaces for improved accessibility for life science researchers. For shotgun metagenomics, such tools have not yet been established, and many analyses can still only be performed via the command-line (i.e., only interacting with a computer via coded commands, with no mouse capacities or graphics to guide the user).

Metagenomics of Meat and Poultry

Limitations associated with use of culture dependent methods have been known in food microbiology for many years. Before NGS became common, food microbiologists attempted to look at unculturable bacteria through other avenues. For example, 16S Ribosomal DNA PCR and Denaturing Gradient Gel Electrophoresis (known as 16S rDNA DGGE fingerprinting) were tools used as a precursor to the modern 16S protocol. The 16S rDNA DGGE separates equal length

DNA fragments based on sequence-specific melting in a polyacrylamide gel with a gradient of a denaturant chemical (Tzeneva et al., 2008). The difference between this method and more modern analysis of 16S is that, after extraction and PCR, fragment lengths are separated on a gel as opposed to being prepared for library sequencing. While widely employed as a precursor to modern day microbial ecology studies, this method did have drawbacks. In addition to less developed whole community DNA extractions, other problems with the method included formation of chimeric sequences (a new artificial gene created when two or more biological genes overlap and combine during PCR) affecting the band distribution, limited DNA fragment length (500 bp), co-migration of DNA fragments (resulting in different fragments having identical melting behavior), and a single species with multiple rRNA copies resulting in multiple bands and diversity overestimates (Ercolini, 2004).

As a result of these issues, shotgun metagenomic techniques are starting to be used to answer specific food safety and quality questions regarding meat and poultry. There are many areas within food microbiology in which metagenomics can aid further research. While targeting pathogenic bacteria is related to food safety, and assessment of specific spoilage bacteria is more associated with meat quality, both safety and quality studies are beginning to look at the entire community of organisms in an environment. In the past, metagenomics in meat and poultry production was primarily used for descriptive purposes; e.g., finding the true polymicrobial diversity of a specific food product or environment (Mayo et al., 2014). As more background understanding is obtained, hypothesis testing is now beginning to emerge, using both 16S and shotgun metagenomics in randomized, controlled studies(Nieminen et al., 2012).

Food Safety

Pathogen mitigation of specific culturable pathogen strains or indicator organisms has been the standard in meat and poultry production as a result of the sensitivity and affordability of these methods. However, some culture methods take up to a few days to confirm a positive presence of pathogens of interest (Valderrama et al., 2016). While rapid diagnostic methods, such as culturing or BAX PCR, allow for fast screening of potential pathogenic bacteria, results are only a presumptive positive or negative and still require cultural confirmation through biochemical, serological, or genetic tests (Hoorfar, 2011). While not currently deployable within the framework of ongoing food safety programs, a future goal of metagenomics is to provide confirmatory rapid pathogen detection in complex sample matrices. Currently, metagenomic research with respect to pathogenic organisms can be broken down into detection and reduction of organisms, as well as more downstream public health application within a greater microbial community.

Pathogenic Bacteria

Detection of Pathogenic Bacteria with Metagenomics. While whole genome sequences for many pathogens are available in the public domain, there has been limited use of metagenomics as a tool in true regulatory pathogen identification. An example of limitations regarding pathogen detection in metagenomic samples was demonstrated in a metagenomic study of a feedlot beef production system (Noyes et al., 2016). This study highlighted both the challenges associated with genetic similarity between many bacterial species and the downfalls of short read analysis. Because shotgun metagenomics has a non-bias enrichment PCR step in library preparation, there is randomness to what portion of any bacterial genome you will sequence. For example, in the beef study, a 100 base pair read assigned to *Salmonella Newport*

reference genome could have come from this serotype; but *Salmonella Typhimurium* and other Enterobacteriaceae, such as *E. coli*, also shared the same conserved genetic region (Yang et al., 2016). As a result of these homologous regions between bacteria of different strains, and even different species, short read sequencing cannot offer a definitive answer to what strain a specific 150 to 250 base-pair fragment belongs to. In situations where strain level differentiation is the difference between an adulterant in a meat product and something considered part of the natural microflora, metagenomic tools are not capable of these distinctions due to intra-species genetic homology and limitations of short-read sequencing.

As a result, pathogen detection within un-enriched samples (i.e., within metagenomic data) is still a future goal and not a current reality. In contrast, whole-genome sequencing (WGS) is already being used in outbreak investigations, as this technique enables highly sensitive differentiation of extremely similar serotypes (even down to one single nucleotide polymorphism or SNP) (Bergholz et al., 2014). The reason that WGS is currently better suited to such applications is its ability to create full-length genomic assemblies from the sequence data, as opposed to metagenomic data, which typically produce very fragmented, low-quality assemblies.

With metagenomic limitations in mind, there has been some work to solve issues associated with pathogen detection in metagenomic data. The creators of SeqSero, a web-based tool for determining *Salmonella* serotypes using high-throughput genome sequencing data from more than 2,300 serotypes, have attempted to identify *Salmonella* in shotgun metagenomic data (Zhang et al., 2015). While the tool is aimed primarily at WGS assembly of *Salmonella*, it does provide the option for pathogen detection within shotgun metagenomic data. The SeqSero developers tested the method by infecting mice with *Salmonella Typhimurium* and then using metagenomic sequencing of DNA extracted from the feces from the mice; they were able to

identify Serotype *Typhimurium*. As a follow up, fecal samples from patients known to harbor *E. coli* O104:H4 also were evaluated using the same tool to test for false positives; no reads from this sample group mapped to *Salmonella* spp. (Zhang et al., 2015). While this work provided a framework for identification of pathogens within metagenomic data, the *Salmonella* in the mice was at an elevated level due to an active illness. By contrast, food associated pathogen detection would need to detect specific strains at a lower relative abundance than those present in clinical infection doses.

Detection of pathogen reduction using metagenomics. While metagenomics was not used during the 2011 European O104:H4 outbreak, retrospective use helped characterize the event. After the outbreak, 45 archived fecal samples from patients were evaluated using shotgun metagenomics. Using this retrospective analysis, researchers were able to create a draft sequence of the outbreak *E. coli* strain, but were only successful at detecting the strain in 67% of the cases that were confirmed via culture (Loman et al., 2013). While not effective at overt pathogen detection, this study did provide evidence of the ability to create draft genomes from metagenomic samples. This demonstrated the usefulness of metagenomics as an approach to characterize outbreak strains and find co-occurrences of outbreaks. Metagenomics also can be paired with other NGS techniques to investigate reductions in pathogen prevalence or abundance. While not directly related to detection of pathogens, a study conducted to identify metabolic pathways of Enterohemorrhagic *E.coli* demonstrated that the pairing of techniques was effective by first using 16S to understand the bacterial community of ground beef, and then looking into metabolic pathways (Galia et al., 2017).

Metagenomics in a Public Health Setting. While there are clear limitations to the clinical application of shotgun metagenomic assessments to detect pathogenic bacteria, there has been

limited use. For example, in a clinical setting (Nakamura et al., 2008), a patient that presented with symptoms consistent with a foodborne illness was tested for presence of several bacterial and viral pathogens within feces through traditional culture methods; however, no pathogen was detected and therefore the illness could not be diagnosed. After the patient recovered from the illness, another stool sample was obtained and the metagenomic DNA was compared to DNA extracted from the original sample collected while the patient was ill. Detection of reads associated with *Campylobacter jejuni* resulted, which then was confirmed by PCR in the 'ill' sample but not the 'recovered' sample (Nakamura et al., 2008). This resulted in the ability to confirm the causative agent of a foodborne illness that was not possible through traditional culture methods. This case demonstrated the utility of shotgun metagenomics in finding a causative pathogen when what pathogen you are looking for beforehand is not clear.

Commensal Bacteria. While identification and quantification of pathogenic bacteria has been the cornerstone of food safety for years, researchers are beginning to understand the importance of commensal bacteria in modulating health and disease of both individuals and environments, such as those introduced via factories or that contaminate meat and poultry products. Bystander bacteria, while not harmful on their own, can be the vessel for carrying genes of interest that escalate risk of disease. Metagenomics allows for increased understanding of relationships between commensal microbes and pathogens because researchers can sequence whole-community DNA to obtain a "full resolution picture" of the microbial community. Currently, identification and quantification of the organisms in a sample are the cornerstone of metagenomic studies, but more advanced analyses are beginning to find their way into the literature, including to answer questions such as what portion of the microbiome is transmissible both intra- and inter-species? What are the mechanisms of this transmission and what—if any—

organisms in the commensal microbial population should be considered more or less 'risky' (Brito and Alm, 2016)?

A specific example of the role that commensal bacteria can play in dissemination of genes of interest lies in the case study of antibiotic resistance genes. When an antibiotic is introduced into a bacterial community, some research has demonstrated that the microbial community can become stressed, resulting in increased rates of horizontal gene transfer (Beaber et al., 2004). Commensal bacteria can donate and accept genes via horizontal gene transfer, and as a result, both the commensal and pathogenic bacteria can acquire, carry, and disseminate these genes (Djordjevic et al., 2013). A widely studied commensal bacterium is generic *E. coli.* While diarrheagenic *E. coli* and enterohemorrhagic or shigatoxin-producing *E.coli* are the main public health concern, understanding harmless *E. coli* aids in the understanding of pathogenic *E. coli*. This is due to the fact that horizontal gene transfer occurs at a higher frequency in bacteria that are most closely related (Soucy et al., 2015).

A specific community dynamic of importance to meat and poultry production is biofilm formation in processing facilities. Shotgun metagenomics has allowed study of multispecies biofilms and their complex microbial interactions, including genetic makeup, metabolite exchange and quorum sensing that may occur between microorganisms in biofilm communities (Giaouris et al., 2015). In biofilm communities, *E. coli* O157:H7 has been shown to possess the ability to co-colonize with commensal bacteria, and shotgun metagenomics assays have been used to study microbial competition for essential macronutrients (Giaouris et al., 2015).

Genes of Interest. Biological hazards in meat and poultry products have primarily been described in the context of a pathogen or toxin of interest (certainly there are exceptions to this, such as *E. coli* differentiation based on virulence). However, with new technologies, instead of

considering the microorganism in its entirety, specific biomarkers can be evaluated (Brul et al., 2012). In this way, pathogenic bacteria can be further delineated based on specific genetic attributes. Two common genetic components of interest in foodborne pathogens are antibiotic resistance genes and virulence factor genes.

Antibiotic Resistance. Antibiotic resistance in meat and poultry products is of concern due to use of antimicrobials in the raising of food producing animals. The apprehension surrounding antibiotic use in livestock stems from concern that consumption of products derived from animals administered antimicrobials could increase risk of antibiotic resistant bacterial infections and treatment failure in humans (World Health Organization, 2012). In the past, antibiotic resistance has mainly been evaluated using cultural methods, such as those of the National Antimicrobial Resistance Monitoring System (The National Antimicrobial Resistance Monitoring System, 2016). This protocol is intended to monitor trends in antimicrobial resistance among enteric bacteria from humans, retail meats, and animals (The National Antimicrobial Resistance Monitoring System, 2016). However, the current workflow calls for enrichment and culturing of specific pathogens of foodborne concern, followed by antimicrobial susceptibility testing of cultured isolates; enrichment changes the relative abundance of organisms in the community and therefore biases results. Furthermore, some of the tested isolates are sequenced as part of the NARMS whole-genome sequencing initiative, after which the sequences are deposited in GenomeTrakr, a public database of genomes (The National Antimicrobial Resistance Monitoring System, 2016). Currently, however, there is no exploration of commensal bacteria in these NARMS samples, and therefore no investigation into the potential for a "commensal" reservoir of antibiotic resistance genes (Marshall et al., 2009). Previous research has demonstrated that many commensal bacteria harbor antibiotic resistance genes (Marshall et

al., 2009). Hence, accessing these commensal bacteria may be critical in furthering our understanding of antimicrobial resistance and its transmission to humans (Marshall et al., 2009). While antimicrobial resistance regulatory and surveillance programs continue to rely on testing of cultured indicator organisms, the research sector has produced studies that identify and quantify all known resistance gene sequences within metagenomic DNA (i.e., the resistome) (Noyes et al., 2016). Such an approach allows for detection of antimicrobial resistance genes within both pathogens and commensal bacteria, and potentially allows an opportunity to characterize potential for horizontal gene transfer events within the microbial population (Crofts et al., 2017).

Another challenge when evaluating antibiotic resistance genes in a metagenomic sample is the dis-uniform methods of analysis. From sampling to processing to data interpretation, there are no standard methods of assessing risk of antibiotic resistant genes in culture-free methods. While some metagenomic studies of antibiotic resistance evaluate the functional assays of resistance, others infer risk based on database (Martínez et al., 2014). As a result of the varying analyses, it is a challenge to directly compare results from one study to the next.

Virulence Factors. Virulence factors have previously been explored by food microbiologists, most notably as a means to differentiate *E. coli* strains, but primarily in the context of pathogen differentiation instead of community-wide characterization. Virulence factors are known to spread through bacterial communities using horizontal gene transfer avenues such as bacteriophages (Penadés et al., 2015). While presence of genes like *eae* and *stx* are used to classify *E. coli* O157:H7 as pathogenic, the ability to identify such genes in a shotgun metagenomic sample is challenging due to the short-read nature of sequences and inability to create robust assemblies (Yang et al., 2016). Furthermore, linkage inferences between genes and

bacteria are limited because library preparation and sequencing do not amplify 'evenly', resulting in uneven sequence coverage. However, understanding what virulence factors are present in a bacterial community can allow researchers to evaluate how these genetic factors interact. An example of this approach was used to conduct a culture-free assessment of natural- versus conventionally-raised beef cattle in North America (Weinroth et al., 2017c).

Environment. In addition to looking at the meat and poultry products themselves, a very appropriate use of shotgun metagenomics is in characterization of processing facilities used to produce meat and poultry products. Analysis of the biogeography of all potential fomites is a discipline known as building ecology (Doyle et al., 2017). In meat or food processing facilities, metagenomic analysis has been used to identify food spoilage bacteria and pathogenic bacteria on processing surfaces (Pothakos et al., 2015). Another valuable capability that metagenomics affords is investigation of co-occurrence of organisms (an example of this is seen in Pothakos et al., (Pothakos et al., 2015) who investigated microbial co-occurrence in a meat manufacturing plant). Such an analysis can be helpful in mitigating pathogens and spoilage organisms because it enables identification of organisms whose abundance is either directly or inversely correlated with the pathogens or spoilage organisms of interest.

Ability to track a microbial community of a given environment is another example of how metagenomics can be used in environmental analysis. In addition to characterizing microbes in a given environmental community, a longitudinal study allows researchers to observe changes in the composition of the microbial community over time, or to compare the characteristics/makeup of a community with those from another environment. For example, in a study conducted on cooked sausage, *Lactobacillus* spp. increased in relative abundance during

production, while *Enterobacteriaceae* declined between the processing steps of meat receiving and final product (Hultman et al., 2015).

Recent advancements in biostatistical capabilities (i.e., Bayesian modeling) also have enhanced source attribution to product or environmental contamination. For example, SourceTracker (Knights et al., 2011) characterizes 16S data, collected from both the location of concern and potential sources of contamination, and attempts to determine which potential source was the actual source of the contamination (e.g., tables in a packing room for ready-to-eat products could be the location of interest and potential sources of contamination could be employee boots, packaging material receiving, and the cooler areas where meat is stored before packaging). While this is could be a valuable tool for improving general understanding of contamination and sanitation activities, the 16S input does not allow for strain level identification, thus limiting the tool's usefulness in identifying a pathogenic contaminant.

The software MetaMLST is a more recent tool that allows for microbial strain tracking and identification within complex shotgun metagenomic data obtained directly from an environmental sample (Zolfo et al., 2017). This tool works by reconstructing multilocus sequence typing (MLST) loci of pathogens (multiple areas of one genome) of interest from publically available databases and comparing them to microbial communities from shotgun metagenomic data (Zolfo et al., 2017). This tool holds promise for identifying pathogen contamination from different areas within a processing facility.

While not directly related to meat and poultry products, metagenomic data also have been used to better understand plant worker health. Understanding which microflora and contaminates that workers might be exposed to throughout their work shifts helps to develop mitigation strategies to prevent possible negative results of exposure. For example, assessment of airborne

bacteria found in the air vents of a slaughter house could become a proxy for worker airborne exposures, including air quality (McLean et al., 2014).

Rearing of Animals. Use of metagenomics is not limited to end product food safety and abattoir sanitation. Many scientists have focused on use of metagenomics to discover potential targets that can be used to influence the microbiome of food animals before slaughter. Objectives of such studies vary, but include reduction of pathogen levels within animals, alteration of growth status, mitigation of unwanted outcomes such as liver abscesses in cattle, and manipulation of the nutrient profile of the meat being produced. As methods become more refined and we have accumulated knowledge of animal microbiomes, research goals have correspondingly evolved. Early studies focused more on characterization of animal microbiomes (Oakley et al., 2014) as a way to understand what organisms were present in normal situations. From there, specific locations within the animal's body that impacted end product food safety were mapped using a microbiome approach (Weinroth et al., 2017a). Metagenomics then were employed to better understand movement of pathogens within livestock populations. For example, metagenomics was used to evaluate Shiga toxin-producing *E. coli* (STEC) in feedlot cattle populations, resulting in the observation that STEC colonization was correlated with a lower diversity of gut microflora, which increased as cattle matured (Mir et al., 2016).

Food Quality

Metagenomics has not just increased our understanding of pathogenic and harmful bacteria, but has also helped us to better understand how all bacteria and organisms change over time, and as a result of using specific antimicrobial sequential interventions in processing plants. For instance, fermentation and aging of meat adds value and longevity to the final products.

While these processes have been characterized through culture-based assays in laboratories, metagenomics allows for these processes to be viewed through an entirely different lens.

Shelf Life. Metagenomics has been used to understand how different processing procedures affect the bacterial community and known causative spoilage organisms. In shelf-life studies, a metagenomics approach was used to characterize the impact of different ingredients, packaging types, and processing variation on shelf life. For example, in marinated broiler meat, the addition of a marinade (which contained, among other things, oil and sugar) altered microbial communities such that spoilage-associated *Leuconostoc gasicomitatum*, *Leuconostoc gelidum* and *Lactobacillus* spp. were favored (Nieminen et al., 2012).

Metagenomic investigation also has been used to understand effects of differing types of packaging on meat storage characteristics. For example, comparison of meat products packaged in air, modified-atmosphere (MAP), under vacuum, and active vacuum packaging revealed that *Brochothrix thermosphacta* comprised a larger proportion of the microbiome of meat packaged in air and MAP during the early days of a 45 shelf life study; while *Pseudomonas* spp. became more dominate in air over time (Ercolini et al., 2011). While many bacterial species present on product had previously been associated with meat spoilage, there also were bacteria found that had not previously been associated with meat spoilage.

In addition to profiling microbes present on meat subjected to shelf-life conditions, a study evaluated whether processing environments and other factors contributed to shelf-life variation. One study of beef steak production investigated the microbiome of steaks, carcass swabs and environmental samples over time (Exploring the Sources of Bacterial Spoilers in Beefsteaks by Culture-Independent High-Throughput Sequencing). The investigators discovered many different types of bacterial contaminants across the samples on day zero of sampling, while

samples from latter time points showed that phyla associated with spoilage outcompeted skin-associated bacterial environmental contaminants (*P. acnes*, *Staph. equorum* and *Staphylococcus sp.*) to become the most dominant phyla in the community. Other metagenomic studies have demonstrated variation in production systems, including lot variation in beef that was packaged in modified atmosphere (Säde et al., 2017).

Fermentation. Culture-based studies associated with fermented meats, such as summer sausage, have been plentiful, but such studies focused on a few organisms and were not reflective of an entire microbiome. A 16S study of Italian salami highlighted the complexity of the ripening process by documenting a diverse bacterial community with higher-than-expected numbers of *Lactobacilli spp.* and *Staphylococci spp.*, as well as presence of previously undocumented rare bacterial families.

Regulatory

Currently, metagenomic methods are primarily used in the research setting (Ercolini, 2013). For a method to be adopted as a benchmark for regulatory compliance, it must be verifiable, sensitive, specific, precise and accurate. Current metagenomic methods do not meet all of these thresholds, but there is clearly regulatory interest in advancing the technology to meet such goals. The current gold standard for monitoring foodborne outbreaks is PulseNet USA. The aim of the network is to rapidly detect multi-state outbreaks caused by foodborne pathogens through PFGE of outbreak samples with subsequent comparison to a national database of PFGE patterns. While still relying on PFGE for some pathogens, PulseNet has begun using WGS for *Listeria monocytogenes* detection, and use in both outbreak investigations and routine surveillance of foodborne pathogens (Deng et al., 2016). Though not currently implemented, use of WGS in PulseNet and other regulatory frameworks is likely the precursor to regulation of meat and

poultry pathogens from a metagenomic perspective (Deng et al., 2016). The ability to trace pathogens, or more specifically, virulence patterns, using metagenomic methods may be the ultimate tool for identifying specific environmental sources of contamination that is of public health concern (Bulut et al., 2017; Yang et al., 2016).

Future Direction

In relation to other techniques applied to the area of meat and poultry microbiology, metagenomics is still an emerging methodology. While many labratories and researchers have demostrated versatility and utility of metegenomics for advancing knowledge within a specific discipline, there still are many areas in which use of metagenomics has yet to be fully explored. In attempting to forecast what metagenomics may mean for meat and poultry science, it may be useful to understand how metagenomics has advanced scientific knowledge within human-oriented fields. Beyond chracterizing different niches in the human body, research was conducted on links between the microbiome and different disease conditions ranging from irritable bowel syndrome to some cancers (Althani et al., 2016). Broad areas within meat and poultry research that could benefit from further development of metagenomic methods include improved usability and accuracy of bioinformatic and statistical techniques, single cell sequencing, and intervention-based studies.

The field of bioinformatics is rapidly expanding, and new techniques are being developed in tandum with rapidly decreasing sequencing costs. An area that has seen growth in the past few years, and that will continue to mature, is development of graphic user interfaces (GUI) that will allow more biologists access to informatic tools. In terms of bioinformatics development, advancements in assembly algorithms specific to metagenomic data are continuing to evolve, such as the release of metaSPAdes (Nurk et al., 2017). There are other non-conventional

methods for bioinformatic development as well, such as crowdsourcing development. Through an open call with associated prize money, crowdsourcing developed 89 new sequence alignment algorithms, 30 more efficient than the previous benchmark, in just two weeks (Lakhani et al., 2013).

Single cell sequencing is another area where growth will likely occur. In shotgun metagenomics, diversity of the microbial community decreases ability to detect strain level variation and specific genes. Multiple displacement amplification (MDA) enables amplification of a single bacterial genome. Use of MDA in conjunction with next-generation sequencing can yield almost complete bacterial genomes (Lasken and McLean, 2014). This method has been used in built environments where bacteria of interest are low in abundance—such as pathogens in hospitals. Use of this method could lead to more complete understanding of low-abundance pathogens and their transmission between the environment and a host or a product (Lasken and McLean, 2014). Because processing facilities deal with these same dynamics (i.e., a built environment with low pathogen presence, such as a further processing ready-to-eat meat plant), the usefulness of this technology is certainly an area worth exploring.

Use of metagenomics will likely continue to transition from primarily exploratory and descriptive studies to causative and intervention based studies. Examples of this within the human literature include studies of *Clostridium difficile* infections and gene editing. In individuals with *Clostridium difficile* infections (CDI), characterized by an over proliferation of *Clostridium difficile*, low volume fecal transplants from healthy donors allow for repopulation of the intestinal microbiome with "healthy" microbes. Recent 16S studies have shown high abundance of *Bacteroidetes* and *Firmicutes* in donor feces, leading to the possibility of CDI treatment with a probiotic specifically enriched for these taxa (Shahinas et al., 2012).

Gene editing, specifically using CRISPR-Cas9 and other systems, hold promise for targeting specific genomes in a metagenomic sample. This technology allows for deletion, insertion, and modification of the DNA sequences of cells which, in turn, allows for control of function of specific genes and regulatory elements (Hsu et al., 2014). This method already is being examined in the context of food safety research. Using the CRISPR-Cas9 system, researchers were able to specifically cleave Shiga toxin genes in bacterial cells, leading to significant reductions in numbers of Shiga toxin-producing *E. coli* (Jia et al., 2017). These authors believe that this system could be further harnessed to decrease foodborne pathogens and other genes of interest in certain microbial communities.

CHAPTER 2

Characterization and Comparison of Market Beef, Dairy and Organic Dairy Cow Resistomes and Microbiomes

Summary

Bacteria on meat that are resistant to antimicrobials, as a result of cattle management practices, is a concern; but there are few data characterizing the nature of this issue for some segments of production. Market cows (mature cows harvested for meat after their usefulness as a dairy cow or beef calf producer has ended) make up a fifth of U.S. beef produced—but little is known about the AMR profile of this sector of beef production. The objectives of this study were to use targeted shotgun metagenomics and characterize the resistome and microbiome of colon content and trimmings from carcasses of non-fed beef, dairy, and organic dairy cows delivered for harvest. Colon content and beef trimmings from two different processing facilities were collected during three visits to each facility—to encompass seasonal variation should it exist. Beta-lactam resistance was found in the highest relative abundance in beef trimmings, while multi-drug resistance was most prevalent in colon content. Beef trimmings resistome was impacted by plant location (but not production method), while the resistome of colon content was not altered by either factor. Beef trimmings resistomes were found to be correlated to microbiome composition, though this was not the case in the colon content. These data provide a baseline characterization of an important segment of the beef industry and will allow a more comprehensive understanding of AMR public health risk.

Introduction

The association between food animal production and antimicrobial resistance (AMR) is a relationship public health officials have expressed interest in (Center for Disease Control and Prevention, 2013). There is a concern that administering antibiotics to food animals could increase AMR bacteria on meat and result in resistant infection in humans for which treatment fails (Economou and Gousia, 2015); a worry both the World Health Organization and the Center for Disease Control and Prevention have voiced (Center for Disease Control and Prevention, 2017a; World Health Organization, 2015). Recently, there has been an increased effort to reduce use of medically important antibiotic in food animal production in the United States through the Veterinary Feed Directive (VFD) (U.S. Federal Registry 80 FR 31707). While AMR is naturally occurring (D'Costa et al., 2011), prevalence within a microbial community can increase through selection pressure (Weinroth et al., 2018c) and horizontal gene transfer (von Wintersdorff et al., 2016).

Antibiotics are used to treat and prevent disease in food producing animals (Boeckel et al., 2015). In both beef and dairy production, antibiotics are used throughout different stages of the animal's lifecycle. In beef cattle, 15.3% of cow-calf operations (businesses that raise calves from birth to weaning) use antibiotics for disease prevention (Sneeringer et al., 2015). On dairy operations, in addition to the administration of antibiotics for diarrhea, respiratory illness, and mastitis, 90.1% of farms administer antibiotics to prevent intra-mammary infections when cows are dry (Sneeringer et al., 2015) as well as to treat infection present at the end of lactation. While the majority of culled market cows are raised in a conventional system that uses antibiotics to improve animal welfare and herd health, organic production (with one requirement for certification being no use of antibiotics) is a small but growing segment of the cattle industry

(USDA-ERS, 2013). One reason for an increase in the organic market share is the perception that organic practices decrease selection pressure for AMR and reduce AMR concerns—though scientific literature on the subject has found not using antibiotics to have little to modest results in reducing AMR (Vikram et al., 2017).

Investigations on how antibiotics alter the resistomes and microbiomes of feedlot cattle and milking dairy cows have employed culture, quantitative PCR (qPCR), or shotgun metagenomics methods (Schmidt et al., 2013; Weinroth et al., 2018c). However, studies of the impact of antibiotic use on AMR in culled cows is sparse (Agga et al., 2016; Chambers et al., 2015; Wichmann et al., 2014). At the same time, market cows and bulls comprise 17 to 19% of the Federally Inspected U.S. slaughter annually (Woerner, 2012); these animals are either culled beef cows or cows culled from dairy operations. Therefore, it is important to gain a baseline understanding of this production sector to accurately understand risk associated with U.S. beef production.

Shotgun metagenomics allows for an ecological survey of AMR genes present in a bacterial community without having to preselect only a few genes of interest. Targeted shotgun metagenomics further refines the ability to look into a community by targeting specific genes and reducing background DNA not of primary interest (Noyes et al., 2017). At the same time, in addition to focusing on the genes that confer resistance, understanding the phylogenetic makeup of the community is vital because some phyla are known to carry AMR genes at a higher frequency than others (Berendonk et al., 2015). Hence, the objective of this study was to use targeted shotgun metagenomic and 16S rRNA amplicon sequencing to characterize the resistome and microbiome of colon content and trimmings from carcasses of beef, dairy, and organic dairy market cows at harvest.

Materials and Methods

Description of Study Population. Cows are defined as bovine females that have been through parturition. In this study, the population was further refined to 'market cows' which are cows that have been in a production system (either dairy or as a producer of offspring for feedlots) but have come to the end of their productive life and are slaughtered for meat. Market cows, in contrast to 'fed-cattle' (cattle that spend the last months of their lives in a feed yard on high concentrate diets) are typically feed higher roughage diets. Also, in contrast to fed-cattle, market cows tend to have a longer lifespan than feedlot cattle, with dairy cows living around 3 to 5 years and beef cows around 8 to 10 years.

Three production systems were sampled in this study: conventional beef cows "CON-B", conventional dairy cows "CON-D", and organic dairy cows "ORG-D." Because of the nature of beef cattle production, there was not a sizable organic market cow beef population to samples, thus this group was excluded from this study. The CON-B cows were composed of cows that produced calves intended for feedlot production, the use of antibiotics was permitted in this production system. The CON-D cows were animals that had produced milk on conventional dairies (where antibiotic usage is permitted). All conventionally managed cattle that are administered antibiotics must wait a minimum amount of time prior to slaughter to assure no antibiotic residue is left in the animal tissue, this is known as a withdrawal time. Finally, ORG-D cows were cows raised on certified organic dairy operations that have several management requirements such as not allowing antibiotic usage, organic feed, and access to the pasture during the growing season.

Processing Facility Overview. Samples were collected over six visits to two U.S. commercial packing facilities that harvested market cows over six months (each plant was

visited three times to help account for seasonal variation). One sampling facility was located in the Southwest, while the other was located in the Midwest. Samples were collected from carcasses of cows that were generated via three production backgrounds at each plant location: CON-B, CON-D, and ORG-D. At each packing facility, 9 composite colon content samples and 9 composite beef trimming samples were collected; three from each production system and sample type (Fig. 2.1).

Colon content sampling. Fifty-four composite colon content samples (each composite was comprised of 9 to 10 individual cow colons composited) were collected in total. At each facility, three composite samples were obtained to reflect each production system, for a total of nine composites per plant per visit (9 samples per visit × 6 visits = 54 samples). Individual colon content samples were acquired by obtaining a sigmoid colon from the evisceration belt, making an incision in the colon, and then transferring approximately 25 g of colon content into a plastic bag; gloves were changed between each colon to prevent cross-contamination. Samples were stored at 4°C and shipped to the USDA Meat Animal Research Center (Clay Center, NE) for further processing. Composite samples were created by combining the 5 to 5.5g from 9 to 10 individual colons (50g total / sample) within treatment group. Composited samples were shipped on ice to Colorado State University (CSU) in Fort Collins, CO and stored at -80°C until further processing.

Trimmings derived from the chilling cooler. Fifty-four composite beef trimmings samples (900 g each) were obtained from carcasses meeting study design requirements while located in the plant chilling cooler 24 h ± 4h after colon content samples were collected. While sampling occurred according to the same experimental design in the cooler, individual carcass identity was not maintained and each sample likely was obtained from different carcasses than those from

which colon content was collected. Approximately 90 to 130 g of bovine trimmings (comprised of a combination of a the brachiocephalicus, trapezius, rhomboideus, and splenius muscles) from each carcass were excised and composited to create three composite samples per production system (7 to 10 carcasses per composite). Samples were immediately placed on ice and transported to CSU for processing.

Trimmings derived from fabrication. Because the plant that was located in the Southwest did not have an organic product label program, it was impossible to maintain identity of carcasses from cows produced in such a fashion through completion of the fabrication (carcass disassembly) process. Therefore, trimmings samples collected from carcasses during fabrication were only collected in the Midwest plant. Furthermore, because conventional beef and dairy cows were not marketed separately in the Midwest plant, the only comparison that could be made was between conventional and organic practices. During each of the three plant visits, three organic and three conventional trim samples were obtained for a total of 18 composites total. Each trimmings sample was approximately 900 g. Samples were immediately put on ice and transported to CSU for processing.

Isolation of DNA for sequencing. Colon content was thawed at 4°C prior to DNA isolation. DNA was isolated from a 0.2 gram aliquot of each composite colon sample using the QIAamp PowerFecal DNA Kit (Qiagen, Hilden, Germany) via manufacturer's instructions.

Upon arrival at CSU, cooler and fabrication trimmings was kept on ice and processed within 24 hours. To each sample bag, 180mL of phosphate-buffered saline (PBS) was added and the bag was hand-massaged. After massaging, all supernatant was centrifuged (10,000 ×g for 10 minutes at 4°C) to pellet intact cells. Pellets were stored at -80°C until DNA isolation. DNA

from each thawed pellet was extracted using the QIAamp PowerFecal DNA Kit (Qiagen, Hilden, Germany) via manufacturer's instructions. DNA was concentrated using an ethanol precipitation.

Library Preparation and Sequencing: shotgun metagenomics. Shotgun metagenomic libraries were prepared using the SureSelectXT-HS Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library (Agilent Technologies) with 'MEGaRICH' (Noyes et al., 2017) a custom-designed target enrichment kit specific to AMR genes, following the protocols described in Noyes et al. (Noyes et al., 2017). Prepared libraries were transported to the UC-Denver Genomics and Microarray Core Facility (Denver, CO), where paired-end (2x150) sequencing was performed on an Illumina NovaSEQ 6000 (Illumina, Inc., San Diego, CA) with a target of 20 million reads per colon sample and 100 million reads per cooler and fabrication trimmings sample. Along with the biological samples, two ZymoBIOMICS mock communities (Zymo Research, Irvine CA) with meta-sequin (Mix A and B) added in at 2% of the mock community DNA by weight (Hardwick et al., 2018) were sequenced on each lane (more information on this can be found in Supplementary methods).

Library Preparation and Sequencing: 16S rRNA Sequencing.. Aliquots of DNA from each sample were shipped to Novogene Corporation (Beijing, China) for 16S rDNA library preparation and sequencing. The V4 region of the 16S subunit was amplified with the 515/806R primer set. Paired-end sequencing (2 x 250) was conducted on an Illumina HiSeq 2500 (Illumina, Inc., San Diego, CA).

Shotgun metagenomic processing. Demultiplexed FASTQ files were transferred from UC Denver. Samples were processed with the AMRplusplus (Lakin et al., 2017) pipeline with modification. Briefly, samples underwent trimming via Trimmomatic (Bolger et al., 2014b) and removal of bovine DNA via the BWA aligner (Li and Durbin, 2009). Duplicate DNA reads

(reads that were identical the entire length of the read), were removed via BBTools' dedupe script (https://jgi.doe.gov/data-and-tools/bbtools/) Once samples were converted from the FASTQ to FASTA format, IDBA-UD (Peng et al., 2012) was used to construct contigs. From there, HMMERv 3.1 (Finn et al., 2015) was used to classify contigs to hidden Markov models (HMM) trained on the MEGARes AMR gene database (Lakin et al., 2017) for identification of AMR genes in the sample. Each model was aggregated into a class, mechanism, and HMM group of AMR genes. Genes that required SNP confirmation to confirm resistance, as defined by the AMRplusplus pipeline (Lakin et al., 2017), were removed from downstream analysis. Count tables were adjusted for lane effect via meta-sequins and a threshold for false-positives for AMR models was set using the mock communities, more information on this can be found in Supplementary methods.

Amplicon processing. Demultiplexed samples were obtained from Novogene and processed with QIIME2 v. 2018.11 (Caporaso et al., 2010a). Files were imported into QIIME2 using the 'qiime tools import' command using the paired end option. Exact sequence variants (ESV) were assigned via DADA2 (Callahan et al., 2016) with the first 20 nucleotides of both the forward and reverse reads trimmed as well as truncation at nucleotide 220 of the forward reads and 230 on the reverse reads. Phylogenetic trees were generated using MAFFT (Katoh and Standley, 2013) v. 7, and FastTree2 (Price et al., 2010). Feature taxonomic classification was conducted using a pretrained Naïve Bayesian classifier trained using the 515/806R primers on the Greengenes database (DeSantis et al., 2006a). Reads that were assigned to chloroplast and mitochondria and those that did not have a kingdom classification were removed. After negative controls were removed, tables were parsed by sample type.

Experimental design. The study was constructed as a 2x3 factorial, where time of collection was treated as random, the first factor was facility (Midwest or Southwest), and the second factor was cow production background (CON-B, CON-D, and ORG-D). Because of uncertainly of the geographically distribution of cattle feeding into each facility, only differences in actual facility not region of harvest could be made. Type 1 error was established at $\alpha = 0.05$, and trends were reported with α at 0.051 to 0.010.

Shotgun metagenomics statistics. Shannon's diversity, used as the measurement for alpha diversity, were evaluated using the 'car' (v. 2.1-6) and 'emmeans' (v. 1.1) packages in R (version 3.4.2) via the 'Anova' and 'lsmeans' functions, respectively. Mean separation was accomplished using the 'pairs' function of the 'emmeans' R package. Non-metric multidimensional scaling (NMDS) ordination was performed using Hellinger transformation and Euclidean distances in the metaMDS function of Vegan (Legendre and Gallagher, 2001), with differences compared using Vegan's 'adonis'. Log_2 fold change were calculated using 'FitZig' function in metagenomeSeq (Paulson et al., 2013a) by fitting multivariate zero-inflated Gaussian mixture models. Limma's 'makecontrast' function (Ritchie et al., 2015) was used for pairwise mean separation, adjusted with the Benjamini-Hochberg procedure (Benjamin and Hochberg, 1995).

Amplicon statistics. Differences in read numbers were assessed with the 'anova' and 'pairs' functions from base R and emmeans, respectively. Tables were rarified as follows: Trimmings derived from the cooler at 48,584 reads, trimmings derived from fabrication at 50,684, and colon content at 92,539 reads; each sample type was rarefied to the lowest number of reads in within a sample type allowing for retention of all samples. Alpha diversity was calculated on a rarified table using Faith's phylogenetic diversity. Beta diversity was assessed with weighted and unweighted UniFrac distances. Differences in alpha diversity were assessed

with a Kruskal-Wallis tests while beta diversity differences were evaluated by the 'adonis' function from the Vegan package v. 2.5-3 for interactions and PERMANOVA for main effects and pairwise comparisons. Differential abundance was calculated with the 'qiime composition' tool using ANCOM (Mandal et al., 2015) on phyla present in more than 1% of the total resistome. UniFrac distances were visualized using principal coordinates analysis plots generated in EMPeror (Vázquez-Baeza et al., 2013).

Results

The interaction between facility and production system the market cows were raised in was evaluated in each model; however, no significant interaction ($P > 0.05$) were observed in any model. As a result, only main effects are presented here as facility and production system independently acted on the resistomes and microbiomes of all sample types.

Targeted Shotgun Metagenomics

A total of 4.8M reads were classified to AMR gene model groups. Colon content, on average, contained more ($P < 0.001$) AMR reads (average 80K, 95% C.I. 65 to 96K) when compared to bovine trimmings derived from the cooler (average 7K, 95% C.I. 0 to 22K) or fabrication (average 5K, 95% C.I. 0 to 31K); though there was no significant difference ($P = 0.99$) between types of trimmings. Within colon content, AMR read number did not differ ($P = 0.252$) by facility; but, between production systems organic cows generated DNA with fewer ($P = 0.009$) AMR reads (average = 43K, 95% C.I. 4 to 81K) than conventional dairy cows (average = 108K, 95% C.I. 70 to 147K). Faculty and production system of beef trimmings derived from the carcass chilling cooler or carcass fabrication did not affect ($P > 0.05$) AMR read number.

The resistome of bovine trimmings derived from carcasses located in the chilling cooler was comprised of hits to resistance to beta-lactams (27%), multiple drugs (22%) and

glycopeptides (14%); tetracycline, macrolide, lincosamides, streptogramines (MLS) and aminoglycoside resistance all accounted for more than 5% of the total resistome individually. The most abundant mechanisms of resistance for samples of beef trimmings obtained in the plant chilling cooler were multi-drug efflux pumps (23%), class A betalactamases (19%) and multi-drug regulators (10%), Fig. 2.2. Trimmings collected from carcasses that underwent fabrication had a higher relative abundance of beta-lactam resistance (55%) followed by between 5 to 10% resistance to aminoglycosides, multi-drug, tetracycline, and glycopeptides. Beef trimmings samples collected during carcass fabrication also had the highest relative abundance of the class A betalactamases (39%) mechanism, followed by multi-drug efflux pumps (6%). Finally, colon content obtained from animals during harvest expressed AMR comprised of multi-drug (42%) and tetracycline (31%) resistance, followed by macrolides, lincosamides, streptogramines (MLS) (7%) and beta-lactam (6%) resistance. At the mechanism level, cattle colon content was comprised of 30% multi-drug efflux pumps, 28% tetracycline resistance ribosomal protection proteins, and multi-drug regulators 10%.

Chilling cooler trimmings resistome differed by facility, not production system. AMR Shannon's diversity was not affected ($P = 0.412$; Fig 2.3A) by production system but was affected ($P = 0.026$) by facility. Beef trimmings obtained from carcasses of cows in the Southwest facility generated greater ($P = 0.003$, Fig 2.3B) Shannon's diversity than beef trimmings derived from carcasses of cows in the Midwest facility. Similarly, overall resistome composition of beef trimmings samples collected from carcasses in chilling coolers were not affected ($P = 0.424$, Fig. 2.3C) by production system, though slaughter facility did influence resistome composition ($P = 0.002$; Fig. 2.3D). Between production programs, at the AMR class level, there were very few differences in the resistome of beef trimmings collected from

carcasses in the chilling coolers: trimethoprim resistance was higher ($P = 0.046$) in organic versus beef and aminoglycoside resistance tended ($P = 0.069$) to be higher in beef than dairy. Facility differences by AMR class were more pronounced: seven classes differed ($P < 0.05$, Fig. 2.4) between plant location, while two additional classes tended to differ ($P = 0.05$ to 0.10, Fig 2.4). Tetracycline resistance was more highly ($P = 0.007$) expressed in chilling cooler trimmings in the Midwest facility and aminoglycoside resistance tended to be ($P = 0.099$) expressed more highly in the Southwest facility; other classes that differed by plant location made up less than 5% of the total resistome individually.

Resistome of conventional and organic fabrication trimmings were similar. No production practice differences ($P < 0.05$) in AMR Shannon's diversity or community composition were detected in beef trimmings derived from carcasses at the end of fabrication. Five classes of resistance differed ($P < 0.05$) between beef trimmings derived from carcasses of cattle raised conventionally vs. organically and obtained during fabrication (rifampin, fusidic acid, metronidazole, aminocoumarins, and bacitracin), but the average expression of these classes was below 1.0 log expression, indicating the log-fold change observed was an artifact of spare counts and not of biological significance.

Colon content resistome overall similar, but many AMR class differed. Shannon's diversity of AMR gene model groups of cow colon content was not affected by facility ($P = 0.982$; Fig 2.5C) or production background ($P = 0.845$; Fig 2.5A). Cow colon content beta diversity was not altered by slaughter facility ($P = 0.803$, Fig 2.5D), though differences in resistome composition for cattle colon content due to production background tended to differ ($P = 0.080$, Fig 2.5B) based on Euclidian distance ordination. While community resistomes of cow colon content resistome was similar, several classes of resistance differed ($P < 0.05$; Fig 2.6) by

production background, though there were no differences ($P > 0.05$) in AMR classes between facilities. There were no AMR class differences ($P > 0.05$) in cow colon content between CON-D and CON-B, and differences between the two conventional treatments when compared to organic followed a similar pattern. Resistance to tetracycline was greater ($P < 0.05$) in colon content of both conventional treatments when compared to ORG-D, as was aminoglycoside resistance in CON-B and beta-lactam resistance in CON-D. Multi-drug resistance tended be higher ($P = 0.069$ CON-D and $P = 0.076$ CON-B) in conventional versus ORG-D. Conversely, relative abundance of rifampicin colon content of ORG-D had higher ($P < 0.05$) expression of rifampin resistance compared to that from conventional cattle, and a greater expression of trimethoprim than colon content from dairy cows. Additionally, aminocoumarin resistance in colon content obtained from ORG-D tended to be greater ($P = 0.069$) than that of colon content of CON-D, while metronidazole and trimethoprim tended to be more highly ($P = 0.076$) expressed in the colon content of ORG-D versus CON-B.

AMR genes decreases in richness through production. Among all sample types collected (colon content, trimming derived from the chilling cooler, and trimmings derived from the fabrication), Shannon's diversity was greater for the resistome in colon contents ($P < 0.001$, Fig 2.7B) than on beef trimmings derived from either production location in the facilities, and the number of gene model groups decreased ($P < 0.05$, Fig 2.7C) the further into production the sample was acquired. Beta diversity by stage of cow processing differed ($P < 0.001$, Fig 2.7A), with a clear separation between colon contents and bovine trimmings, though bovine trimmings derived from the two different locations in the plants were similar to each other.

Amplicon Sequencing Results

After quality filtering, 17.5M reads were retained for microbiome analysis (average = 139K, range 54K to 194K). The number of reads generated per sample type differed ($P < 0$ .0001) as colon content had a greater ($P < 0.001$) numbers of reads compared to both types of bovine trimmings. Colon samples generated an average of 171K (95% C. I. 163 to 178K) filtered reads per sample, bovine trimmings from the chilling cooler 118K reads (95% C. I. 110 to 125K), and bovine trimmings from fabrication 108K reads (range 96 to 120K) on average per sample.

Relative abundance of Proteobacteria (63%), Firmicutes (19%) and Bacteroidetes (7%) were greatest in beef trimmings derived from cow carcasses in the chilling cooler, Fig. 2.8. Actinobacteria and Acidobacteria individually made up 5 and 1% of colon content microbiome, respectively. Similar to trimmings derived from carcasses in the chilling cooler, trimmings derived from carcasses during fabrication were mainly comprised of Proteobacteria (49%), Firmicutes (25%) and Bacteroidetes (10%), Fig 2.8. Four other phyla, Actinobacteria, Acidobacteria, Gemmatimonadetes, and Chloroflexi made up 1 to 7% of the total microbiome of beef trimmings obtained from carcasses during fabrication. In contrast to the beef trimmings samples, colon content was comprised primarily of Firmicutes (60%) followed by Bacteroidetes (27%). Proteobacteria, Tenericutes, Verrucomicrobia, Actinobacteria and Spirochaetes made up between 1 to 5% of the total microbiome in cow colon samples, Fig. 2.8.

Facilities had different cooler trimming colon content Microbiomes

Alpha diversity of beef trimmings collected from carcasses in the chilling coolers by did not differ by production background ($P = 0.98$) or plant location ($P = 0.28$). Beta diversity of trimmings derived from the cooler differed by facility (weighted UniFrac $P = 0.005$, unweighted

UniFrac $P = 0.646$; Fig. 2.9) but not by the production system the cows were raised in (weighted UniFrac $P = 0.646$; unweighted UniFrac $P = 0.874$; Fig. 2.9) When log-fold change was assessed at the phylum level between facilities of slaughter, no differences ($P > 0.05$) were detected at the phyla level in trimmings derived from the cooler.

Alpha diversity of colon content was different ($P = 0.011$) between facilities but not production practices ($P = 0.397$). Colon content beta diversity differed (weighted UniFrac $P = 0.001$; unweighted UniFrac $P = 0.001$; Fig. 2.9) by facility, while the production system the cows were raised in did not (weighted UniFrac $P = 0.092$, unweighted UniFrac $P = 0.224$; Fig. 2.9), though weighted UniFrac differences between production systems were considered a trend. At the phyla level, no differences ($P > 0.05$) among treatments in microbiome were observed in colon content. Between facilities, cow colon content in the Southwest processing plant had a greater amount ($P < 0.05$; $W = 46$) of the phyla Fusobacteria than colon content from samples collected in the Midwest facility. Upon further investigation, Fusobacteria was present in 48% of cow colon content collected from the Midwest facility and 70% of colon content samples from the Southwest facility. Average total relative abundance in colon samples of Fusobacteria was 0.01% in the Midwest facility (range 0.00 to 0.15%) and 0.15% in the Southwest facility (range 0.00 to 2.41%).

In the Midwest Facility, trimming microbiome was affected by season and production system

In trimmings obtained from carcasses during fabrication, trimmings from conventional cattle had higher alpha diversity ($P = 0.046$; Fig. 2.10) and the two types of trimmings had differences (weighted UniFrac $P = 0.003$, unweighted UniFrac $P = 0.067$; Fig 2.10) in beta diversity. However, there were no phyla level differences ($P > 0.05$) between conventional and organic trimmings derived from fabrication. Between trimmings from the Midwest facility

collected in the chilling cooler and fabrication, alpha and beta diversity did not differ (alpha P = 0.47; weighted UniFrac P = 0.37, unweighted UniFrac P = 0.552; Fig. 2.11). However, season was seen to be a significant (weighted UniFrac P = 0.001; unweighted UniFrac P = 0.001; Fig. 2.11) effect the composition of the trimmings microbiome.

Discussion

These data provided a comprehensive look into AMR proliferation and differences in market cows colon content and bovine trimmings and the associated microbiomes, Cow colon content had a distinct resistome and microbiome when compared to beef trimmings derived from carcasses in the chilling cooler, and both production system and facility location provided information about the resistome and microbiome of colon content and trimmings derived from both the chilling cooler and fabrication.

Resistomes were different within sample types across all production background groups and facility locations. Beef trimmings derived from carcasses in the chilling cooler had high proportions of resistance to beta-lactams, multi-drug resistance, and glycopeptides. Trimmings derived from carcasses during fabrication also had the greatest number of resistance genes assigned to beta-lactam resistance, but in a higher relative abundance than those obtained from carcasses in the chilling cooler (55% vs. 27%). Other prominent classes of AMR hits for bovine trimmings collected during fabrication were assigned to aminoglycosides, multi-drug, tetracycline, and glycopeptides. When the same targeted metagenomic method was used to investigate bovine trimmings (with different bioinformatic processing), another bovine trimmings study found beta-lactams to be the most abundant class of resistance (Weinroth et al., 2018b), while a study on ground beef found that tetracycline was the major class of resistance (Thomas et al., 2018). It is likely that these study difference could be a result of overall sparse

numbers of resistance genes in end product, thus small changes in actual AMR genes could dramatically alter the relative abundance of different classes. Additionally, the current study deduplicated reads during processing (the removal of reads that are identical within a samples as they are thought to be a result of PCR bias), while the two aforementioned studies did not, which could also explain differences in relative abundance. Culture methods also have detected high levels of tetracycline resistance, with a study of *Enterococci* from retail ground beef finding tetracycline to be the most common class of resistance (Tyson et al., 2018). Resistance to the beta-lactam class of antibiotics has been found in some *Salmonella* serotype isolates (White et al., 2001), though the overall prevalence was more challenging to measure as many studies focus on subsets of beta-lactam resistance that are a higher generation.

Colon content resistome, used to represent the effect of different production strategies on the cows, also was characterized in the present study. The cow colon resistome was more robust than the trimmings resistome, with significantly more reads from colon content attributed to AMR. Colon content was comprised mainly of multi-drug and tetracycline resistance. Tetracycline resistance is very common in cattle feces and has been documented in dairy (Wichmann et al., 2014) and feedlot cattle (Weinroth et al., 2018c), as well as in colon content (Vikram et al., 2017). Contrary to a high level of multi-drug resistance found in this study, other studies have reported low multi-drug resistance in the feces (Noyes et al., 2016) when employing shotgun metagenomics; though multi-drug resistance in the content of pathogens such as *Salmonella* spp. has been reported on the hides of beef cattle (Brichta-Harhay et al., 2011) at a higher prevalence. It is unclear if the reason behind this level of multi-drug resistance was a result of different cattle management practices (market cows have longer life spans than those of

conventional beef fed cattle) or bioinformatic differences (the differing classification of what constitutes multi-drug resistance within different databases and assays).

Once resistomes were characterized within sample type, differences associated with common production management methods and facility locations of harvest were evaluated. In both sample types, across all metrics measured (diversity, community composition, and class level $\log_2$ fold change), there were no interactions between region and treatment group, meaning that these two factors acted independently on the resistome.

The resistome of beef trimmings obtained from cow carcasses in the chilling cooler were not affected by production background (this result was reinforced in trimmings obtained from carcasses during fabrication when conventional and organic samples were compared), though plant location did have an impact on diversity and community composition. When evaluated at the class level, several classes of resistance differed by plant location. This result was explained in the context of the National Antimicrobial Monitoring System, in which data for ground beef AMR across the county (National Antimicrobial Resistance Monitoring System, 2017) illustrates diversity of AMR genes of different classes and culture organisms throughout the country. More broadly, differences in AMR community structure have been noted in soils in different parts of the county, indicating that AMR genes differ by geographical location—and not just within cattle, but also the environment (Forsberg et al., 2014).

In contrast to findings associated with beef trimmings samples, colon content resistome beta diversity trended to be different by production background, but not facility location, this could indicate that trimmings are more influenced by facility environmental factors while colon content could be more influenced by the direct application of antimicrobials. Additionally, many classes of resistance differed among treatment groups: specifically, colon content samples from

organic cattle were comprised of more rifampin resistance genes and fewer tetracycline resistance genes compared to both conventional dairy or beef production background groups. Vikram and coauthors (Vikram et al., 2017)evaluated colon content in feedlot cattle and found that samples from conventional cattle had more tetracycline resistance when evaluated by qPCR, though when tetracycline was evaluated using metagenomics, they found no difference between production system. Using qPRC, the same feedlot study (Vikram et al., 2017) found that some genes associated with MLS and aminoglycoside resistance comprised a higher proportion of all AMR genes in conventional cattle, but no difference among samples from groups in beta-lactam resistance genes. The present study also identified a higher proportion of aminoglycoside resistance genes (colon content samples from beef, and a tendency to be greater in dairy cows, when compared to organic), but differed from the qPCR study in respect to MLS and beta-lactam resistance, as this study found no difference in the MLS class but an increase in beta-lactam resistance. Interestingly, other metagenomic studies (Noyes et al., 2016; Weinroth et al., 2018c) of feedlot cattle reported increased relative abundance of aminoglycoside resistance genes in feces over time during feeding, even though neither study had administered that class of antibiotics. This shift in aminoglycoside resistance, not tied to administration of such drugs during production, may indicate that aminoglycoside resistance is driven by factors outside of direct selection pressure and may be a result of co-selection with other resistance classes.

Differing effects of sample type on results due to the production system the cows were raised in and region is not unexpected. While colon content was used as a metric to assess animal and environmental impact of different regions of slaughter and production practices, trimmings were used as the closest product to distribution, and thus as a potential indicator of human public health risk. In fact, it has been reported that there is not a strong correlation between the fecal

and meat resistomes (Weinroth et al., 2018a). As such, while colon content undoubtedly is more indicative of the changes to the animal as a result of treatments, trimmings may be a better gauge of processing facility differences than production system.

Across sample types, there were no substantial community resistome differences that were associated with production background. While past work has concluded that treatment with antibiotics can result in increased abundance of resistance genes (Chambers et al., 2015), other studies have shown this effect to be proportional to the dosage and duration of antibiotic administration and that there is not always a permanent effect on the resistome (Kanwar et al., 2014; Weinroth et al., 2018c). Although no specific drug application records were collected in this study, aside from overall labeling differences in production background, it was likely that these cattle were never provided diets containing antibiotics, which may explain the lack of substantial separation between treatment groups.

The collection of samples at two different sampling points in the midwestern facility (trimmings from the chilling cooler and trimmings from the end of fabrication) allowed for several comparisons specific to the Midwest facility. First, it was found that organic cows had a different, less diverse microbiome than conventional cows at the end of fabrication. However, this result is confounded different production systems are run in the facility. More specifically, organic cows were always the very first cows to go through fabrication after the facility was cleaned for the day. After the organic cattle were fabricated, while trimmings were removed from food contact surfaces, there was not another cleaning. As a result, it is likely there was a greater number of bacteria present at the start of the conventional cattle. The likelihood that these differences are a result of facility sanitation practices and not a biological difference is reinforced by the fact that there were not differences in the production system the cows were raised in in the

trimmings derived from the chilling cooler. In fact, when trimmings from the Midwest facility were visualized on a principle coordinate analysis, seasonality was seen to be a bigger driver of separation than production system or stage of production.

The microbiome also was investigated, as several studies have identified a link between the composition of the resistome and the composition (taxa-wise) of the microbiome (Hu et al., 2016; Stokes and Gillings, 2011). As with resistome composition, facility of harvest and the production system the cows were raised in acted independently on microbiome composition. Within both colon content and trimmings derived from the chilling cooler, production system the cattle were raised in did not affect the microbiome, through the facility of slaughter did in regard to beta diversity. As with the beef trimmings resistomes, the differences in UniFrac distances between facilities may have been more indicative of processing facility differences than production background influences. After carcass hides were removed during harvesting, they were exposed to several environmental factors unique to each of the different processing facilities. For example, an assessment of airborne bacteria in a harvest facility found that air vents contained actinobacteria, firmicutes and proteobacteria as the dominate phyla present.

Also, of note within the colon content microbiome between regions, was a greater number of samples with a greater mean relative abundance of Fusobacteriaceae of colon content from cattle harvested in the Southwest that contained when compared to those harvested in the Midwest (and at an average higher rate as well). The family Fusobacteriaceae contains *Fusobacterium necrophorum,* which has classically been implicated as the causative organism of liver abscessation (Nagaraja and Lechtenberg, 2007).

On each of the 6 visits to the packing facilities the following samples were collected:

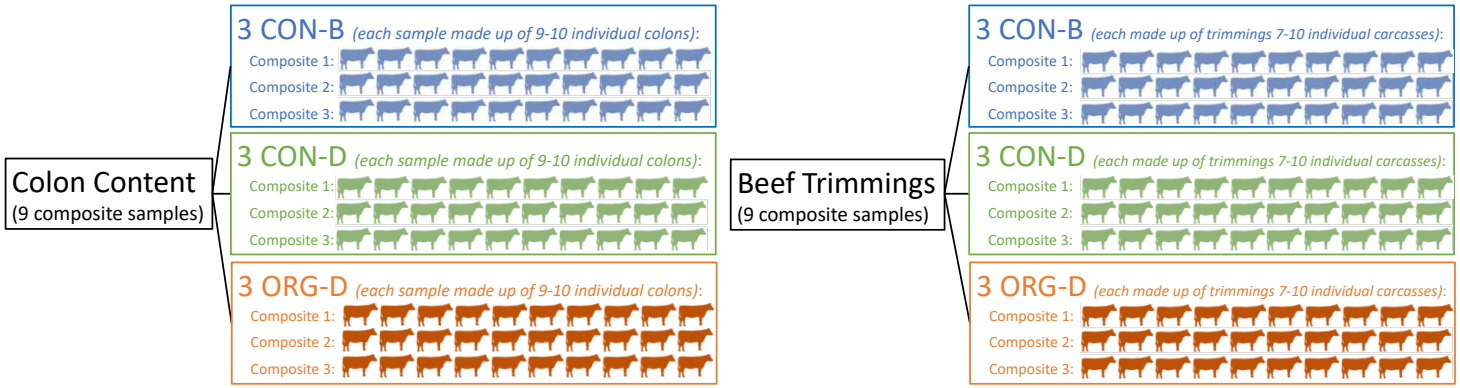


Figure 2.1: Layout of experimental design for each facility visit; this samples scheme was repeated six times over two facilities.

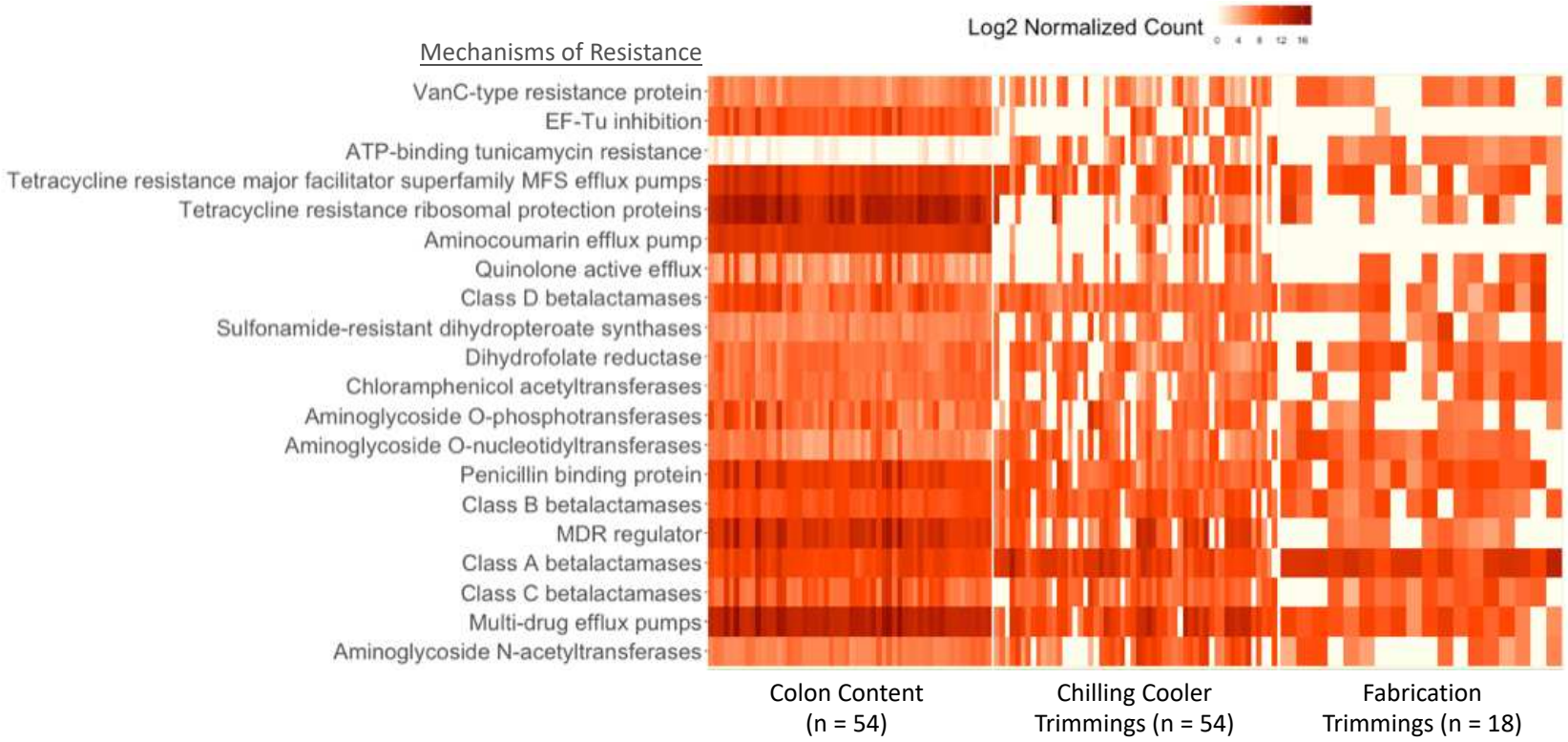
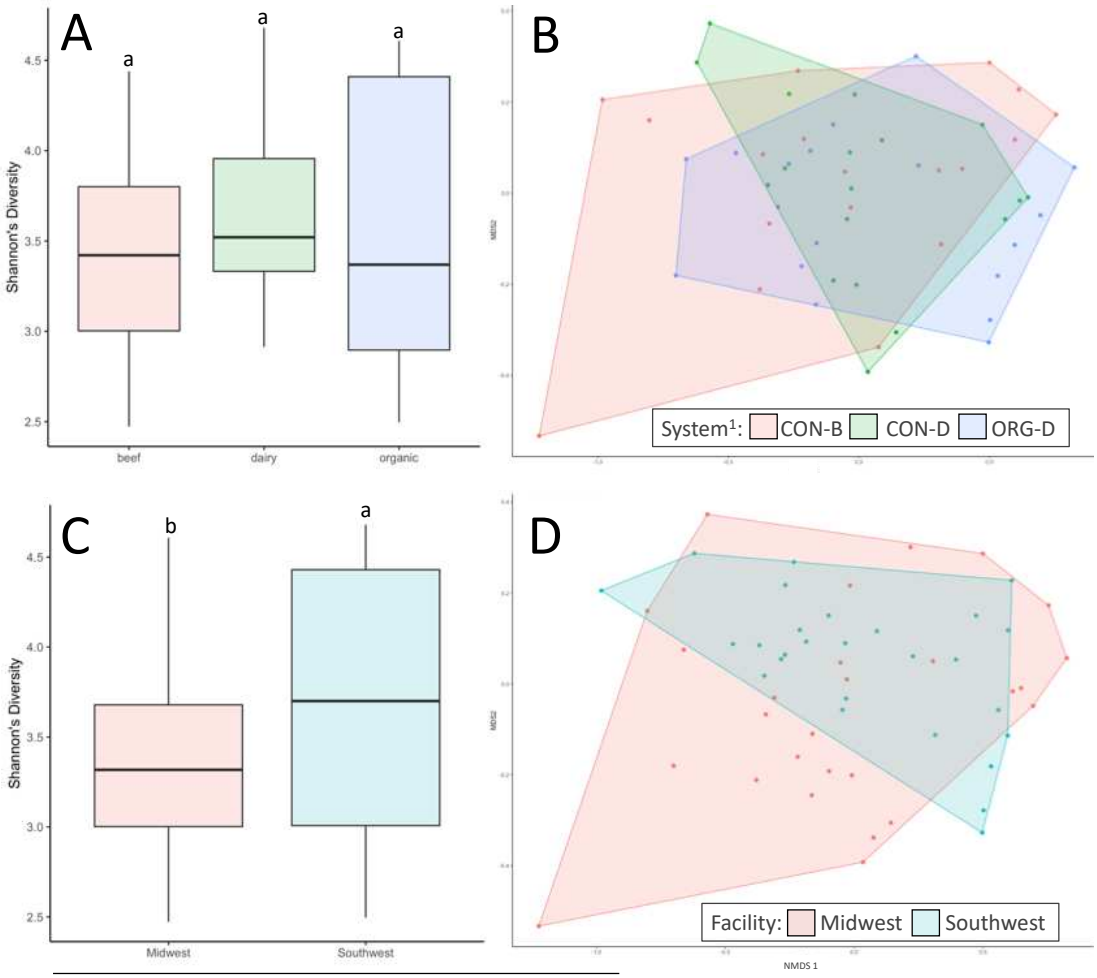


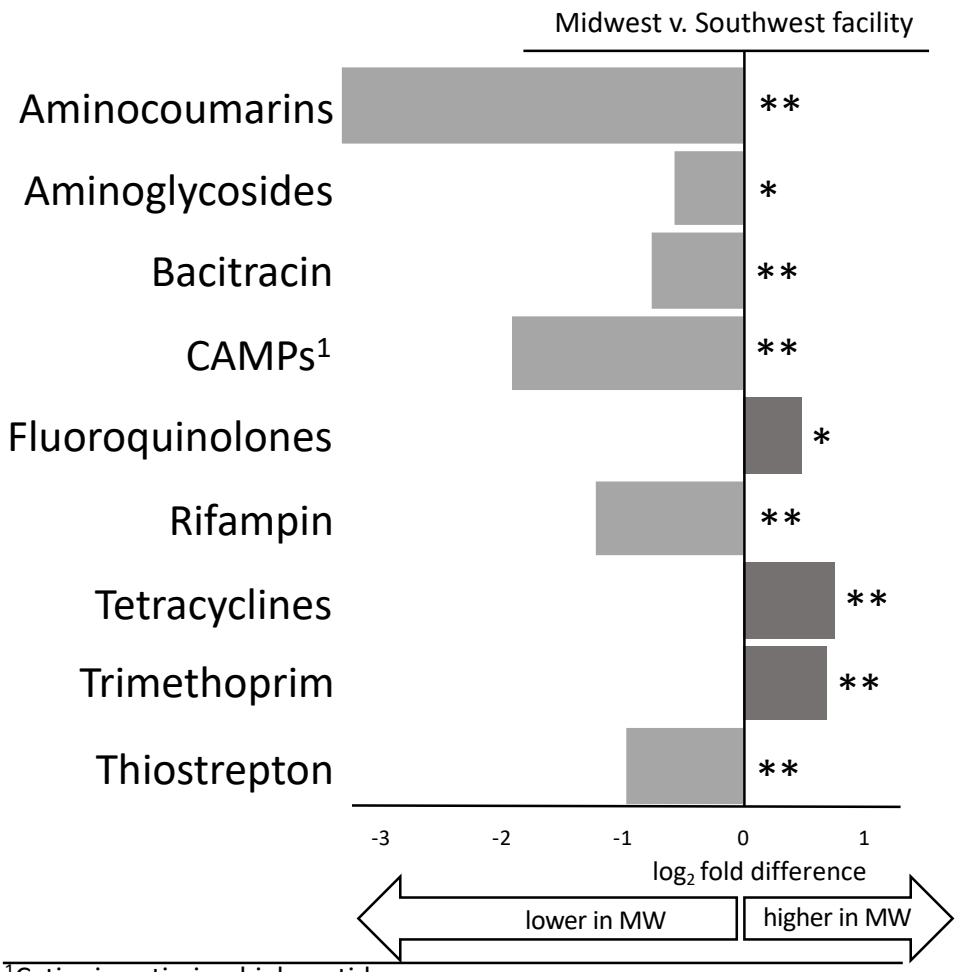
Figure 2.2: Log_2 abundance of the 20 highest abundance mechanisms of resistance present in samples divided by samples type.



[ab] Boxes that bear a different superscript after different ($P < 0.05$)

[1] Production System: conventional beef cows "CON-B", conventional dairy cows "CON-D", and organic dairy cows "ORG-D"

Figure 2.3: Diversity and ordination of resistome trimmings samples derived from the chilling cooler by treatment and region. Shannon's diversity did not differ ($P = 0.0412$) by production system (A) but between facilities, the Southwest facility had higher ($P = 0.003$) Shannon's diversity when compared to the Midwest facility (C). Non-metric multidimensional scaling of Euclidean distance revealed no difference between production practices ($P = 0.406$, R2 = 0.04; B), though facility of harvest did differ ($P = 0.002$, R2 = 0.07, D).



[1]Cationic antimicrobial peptides

**Within a comparison, classes with two asterisks are different ($P < 0.05$)

*Within a comparison, classes with an asterisk are considered a trend ($P = 0.05$ to 0.10)

Figure 2.4: Log2 Fold change of antibiotic resistance classes (that were present in more the 0.1% of the total resistome) of chilling cooler derived trimmings across treatment types between facilities of harvest. Facility of harvest was found to have an effect ($P < 0.05$) on the expression of several classes of resistance. Bars to the left of the midline indicate a lower expression in the Midwest while bars to the right indicate a higher expression in the Midwest.

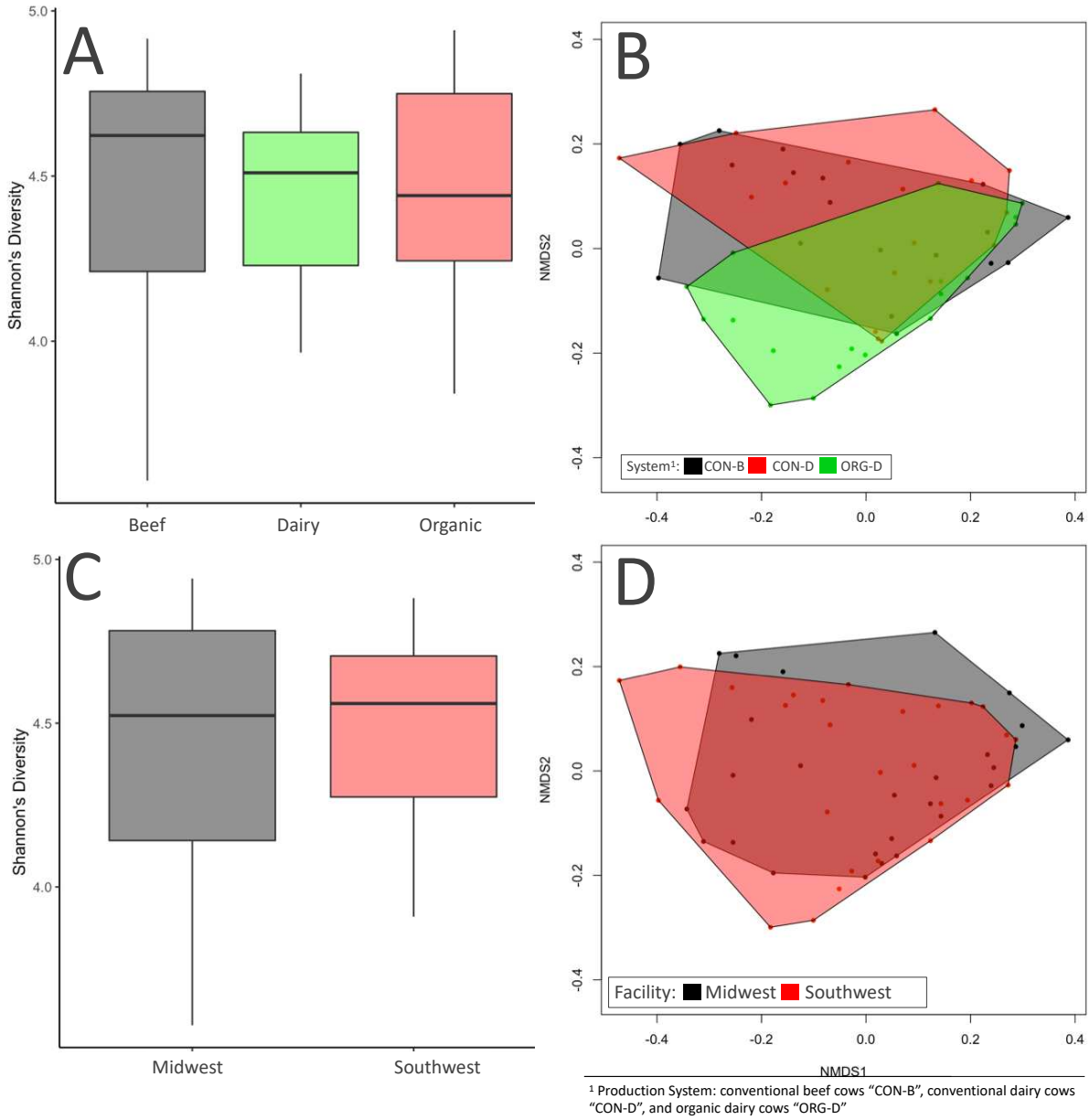
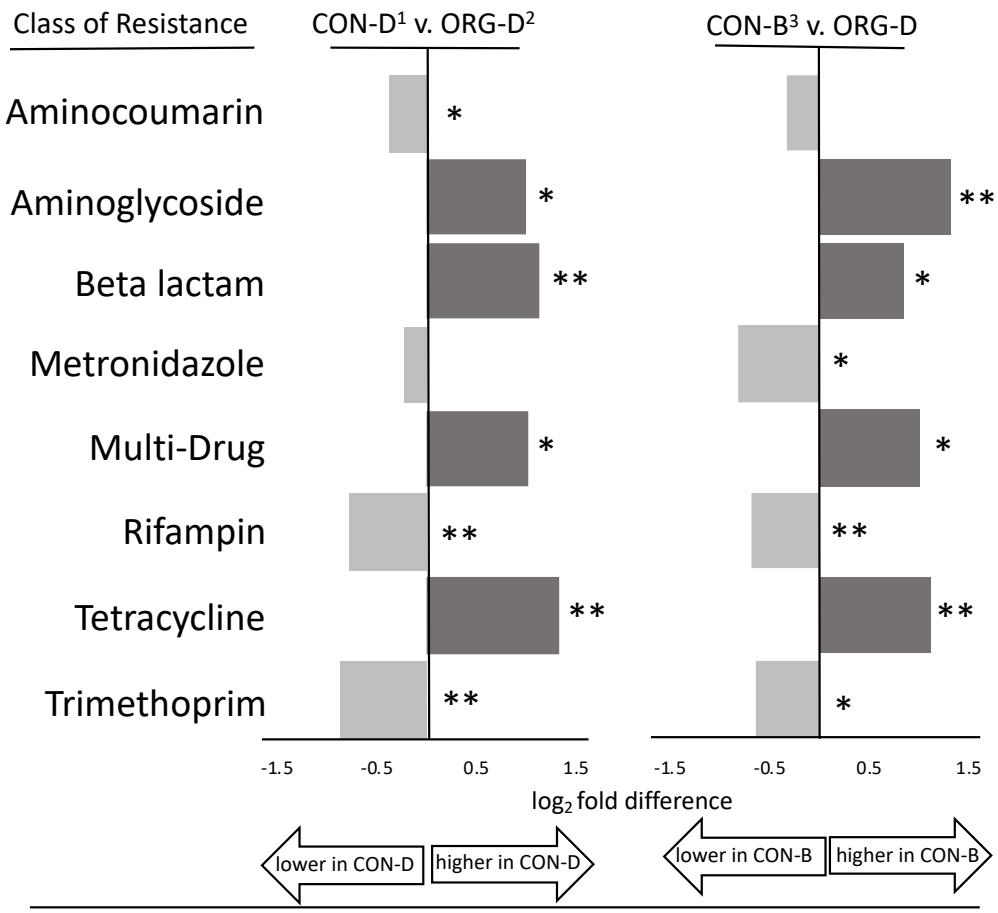


Figure 2.5: Diversity and ordination of the colon content resistome by production system and facility of harvest. Shannon's diversity did not differ ($P < 0.05$) between production systems (A) or facility of harvest (C). Non-metric multidimensional scaling of Euclidean distance revealed no differences between facility of harvest ($P = 0.803$, $R^2 = 0.00$), though production system tended to differ ($P = 0.080$, $R^2 = 0.08$).



**Within a comparison, classes with two asterisks are different ($P < 0.05$)

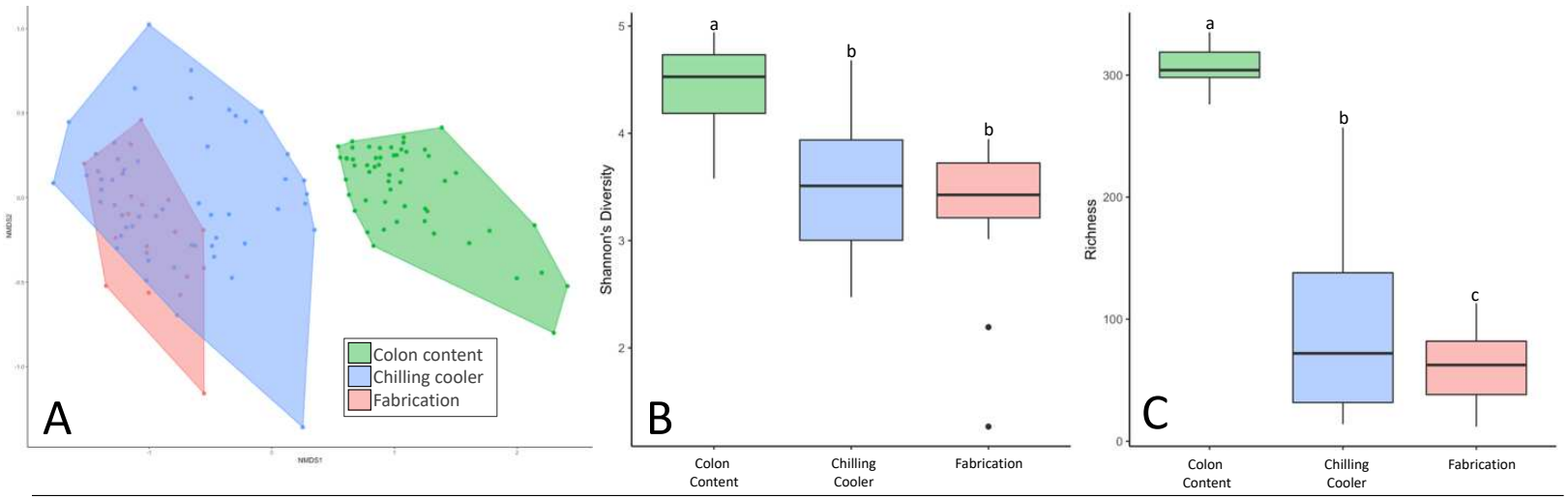
*Within a comparison, classes with an asterisk are considered a trend ($P = 0.05$ to 0.10)

[1] Conventionally raised dairy cows

[2] Organically raised dairy cows

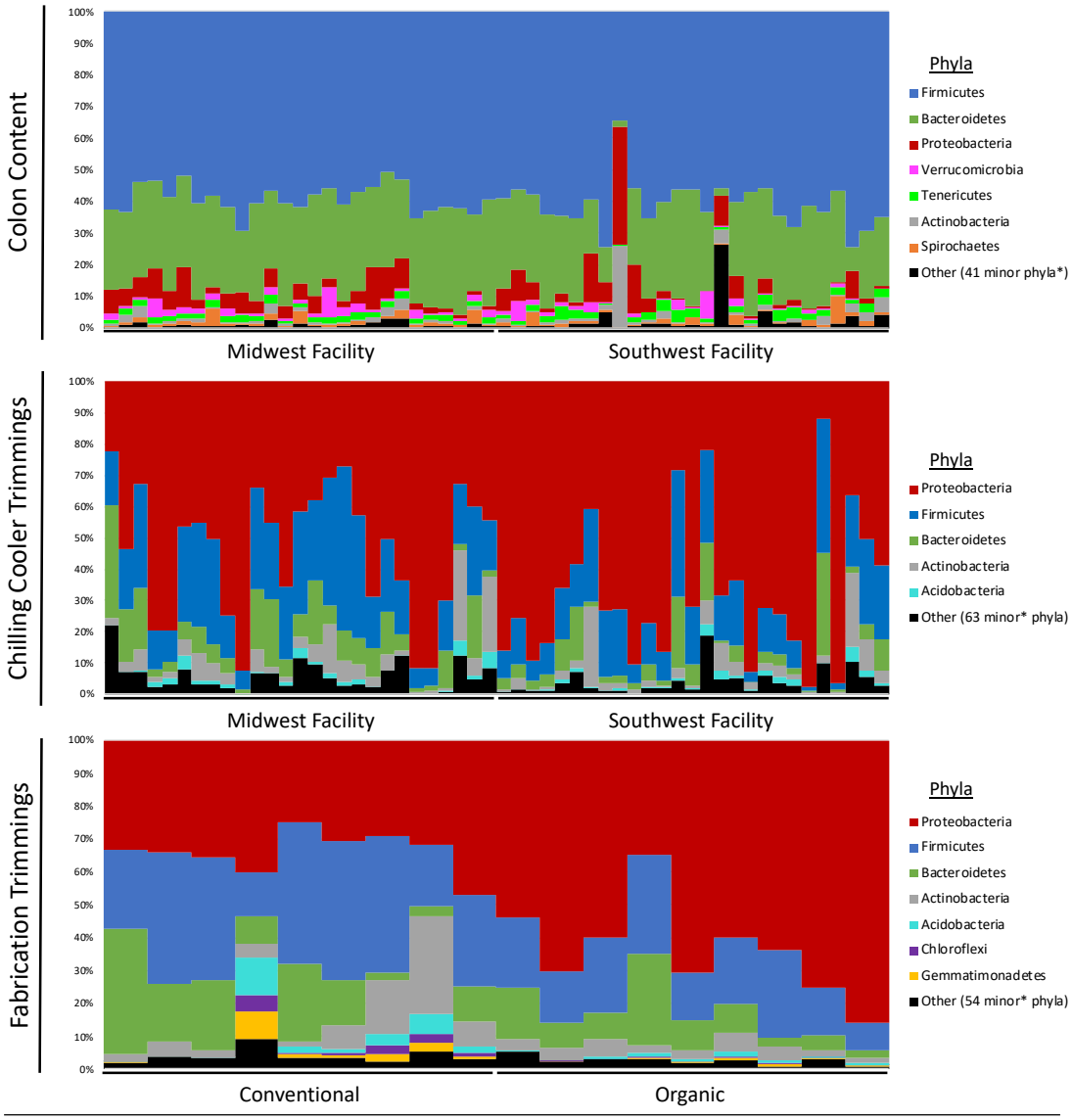
[3] Conventionally raised beef-type cows

Figure 2.6: Log_2 Fold change of classes of antibiotic resistance (that were present in more the 0.1% of the total resistome) of colon content across facilities between production system. Overall resistome tended ($P = 0.08$) to differ between production systems and several classes of resistance were different ($P < 0.05$) between both types of conventional samples and organic; though there were no differences ($P > 0.05$) between conventional beef and dairy colon content by class. Within each panel, bars to the left of the midline indicate a lower expression in the dairy or beef when compared to organic colon content while bars to the right indicate a higher expression in beef or dairy cattle when compared to organic cattle.



[abc] Within panel, boxes that bear different superscripts are different ($P < 0.05$)

Figure 2.7: Resistome differences between stages of production (colon content, trimmings derived from the chilling cooler and trimmings derived from fabrication). Stages of production differed ($P < 0.001$) at the community level and differences ($P < 0.05$) in Shannon's diversity and richness were also observed.



*Other phyla are composed of phyla that were identified in the sample group but not present in more than 1.0% of the total microbiome.

Figure 2.8: Taxonomic composition at the phylum level of (top) colon content, (middle) trimmings derived from the chilling cooler, (bottom) trimmings derived from fabrication.

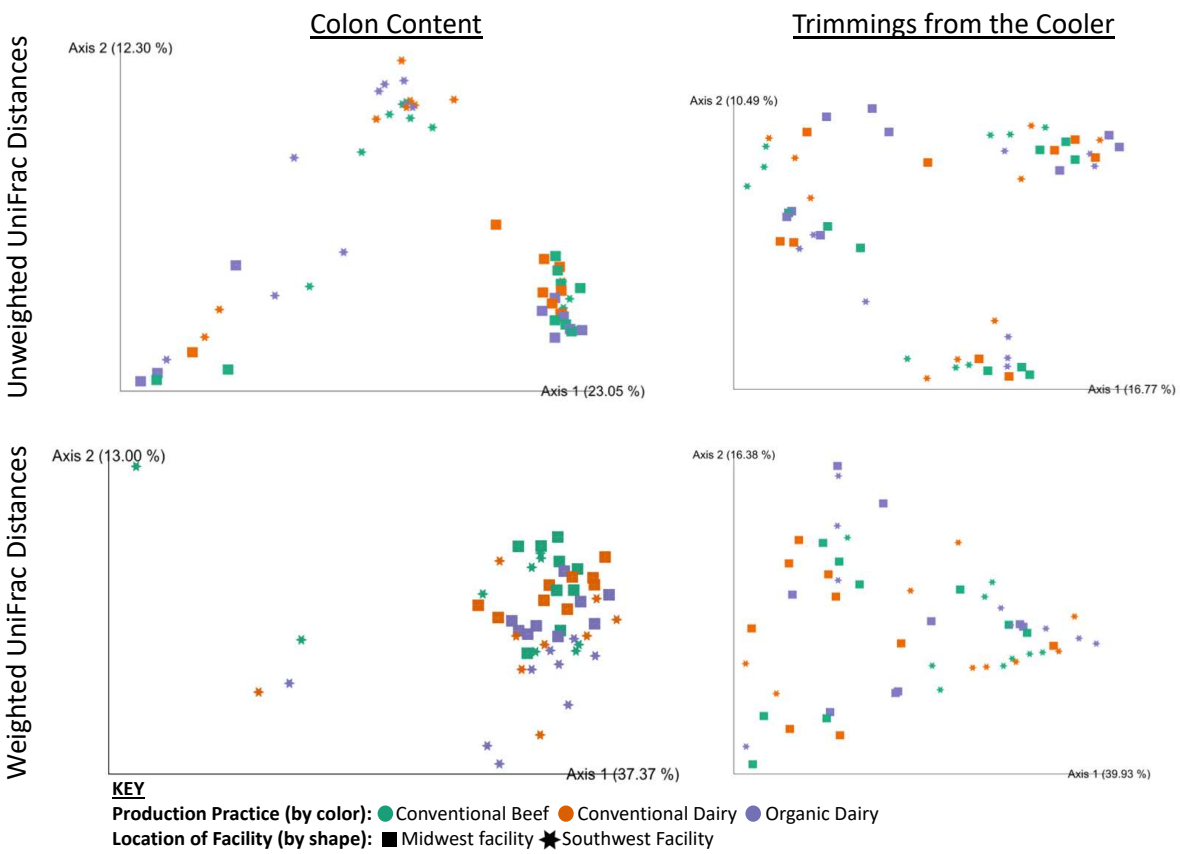


Figure 2.9: Beta diversity of colon content and trimmings derived from the chilling cooler (as measured by weighted and unweighted UniFrac. Samples are colored by production practice the cows were raised in and the shape corresponds to the facility the cows were slaughtered in. In trimmings derived from the cooler, both weighted and unweighted UniFrac distances were different ($P < 0.05$) between facility of harvest but not the production facility the cows were raised in ($P > 0.05$). In colon content, facility of harvest had different colon content ($P < 0.05$) but not production system weighted UniFrac distances ($P = 0.224$); production system unweighted UniFrac distance differences; though weighted UniFrac differences between production systems were considered a trend ($P = 0.092$).

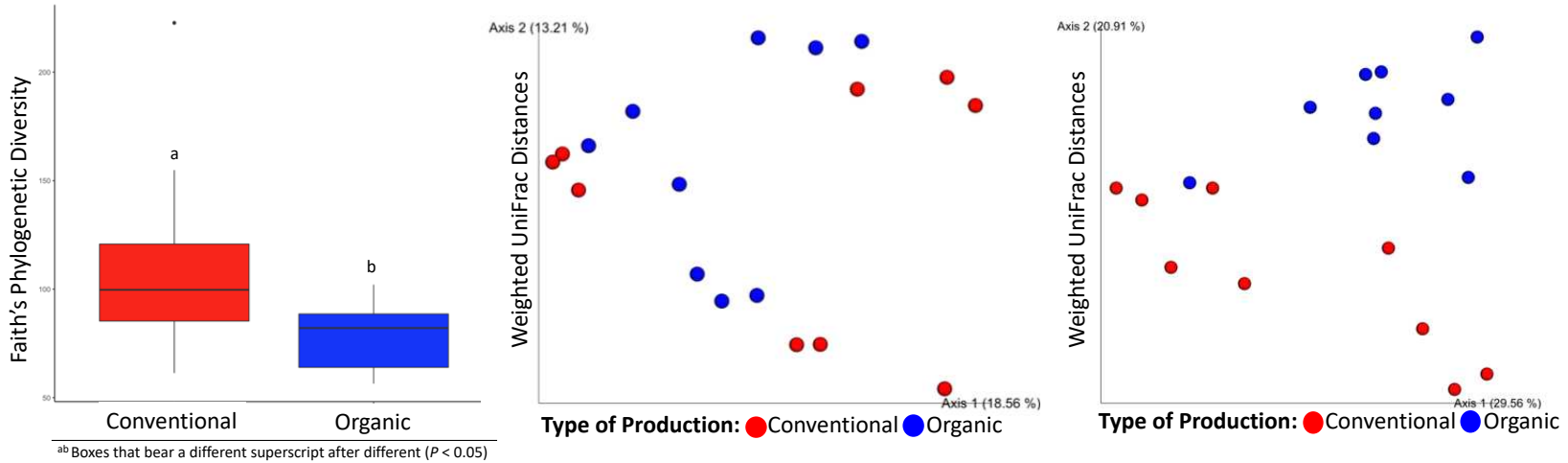


Figure 2.10: Trimmings derived from the end of fabrication in the Midwestern facility colored by production system the cows were raised in. Alpha diversity differed (P < 0.05) between the two production groups as conventional cattle has a higher diversity. Beta diversity also differed (P < 0.05) between conventionally and organically raised cows.

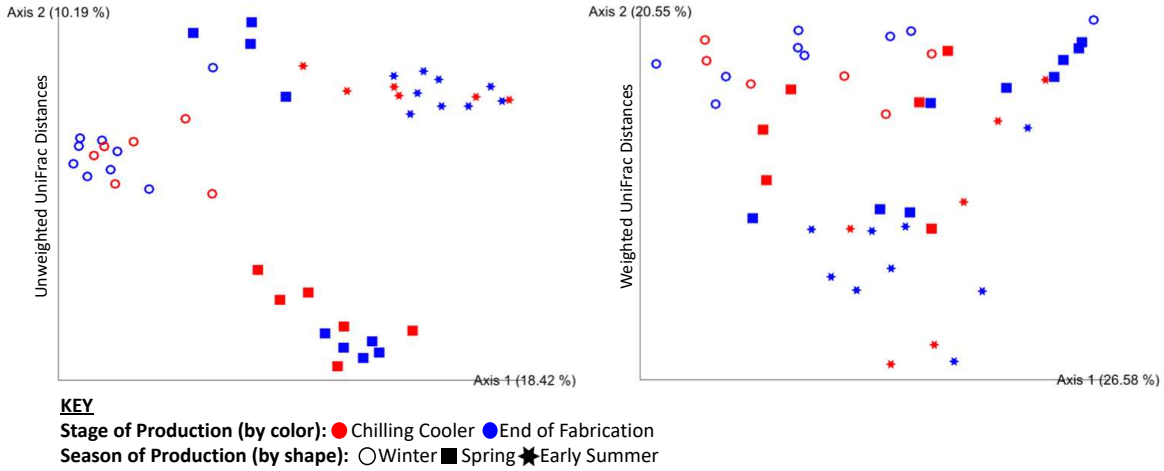


Figure 2.11: Comparison of beta diversity (both weighted and unweighted UniFrac distances) of trimmings derived from different stages of production in the midwestern facility. The samples are colored by stage of production and shaped by season of production. While stage of production did not account for a significant ($P < 0.05$) amount of variation, though seasonality did ($P > 0.05$).

CHAPTER 3

Effect of Tylosin Exposure and Geographical Region of Production on Liver Abscess Rates, Microbiomes, and Resistomes in North American Feedlot Cattle

Summary

Liver abscesses in feedlot cattle are detrimental to animal performance and economic return. Tylosin, a macrolide antibiotic, is used to reduce prevalence of liver abscesses; though it does not always reduce them with the same efficacy. Additionally, with growing concern over antibiotic resistance, there has been increased scrutiny in regard to using antibiotics in food animal production. The objective of this study was to characterize microbiome and resistome differences among cattle administered or not administered tylosin across different feedlot locations with differing liver abscess rates. Cattle (total of 2,256) from three geographical regions were either fed or not fed tylosin. Feces and pen soil samples were collected before harvest and liver abscesses were scored at harvest. Shotgun metagenomics and 16S rRNA amplicon sequencing, as well as culture samples (feces only), were evaluated. The microbiomes and resistomes of cattle did not differ ($P > 0.05$) as a result of tylosin treatment. However, feedlot location did have an effect ($P \leq 0.05$) on cattle's resistomes and microbiomes. Using LASSO, a predictive model was constructed that selected four fecal phyla (Euryarchaeota, Fibrobacteres, candidate phyla Cloacimonetes [WWE1], and WPS2) and two soil phyla (Deferribacteres and Firmicutes) to predict liver abscess rates. This model explained 75% of the variation in liver abscess rates, though a larger sample size is needed to increase universal robustness of the model. These data suggest that tylosin exposure does not have a meaningful impact on overall antibiotic community makeup or microbiome; but that location of cattle

production is a driver of both the community of resistance genes presences and microbiome composition.

Introduction

Liver abscesses in feedlot cattle negatively impact animal efficiency and economic gain, in both the feedlot and at the subsequent packing plant. Liver abscess rates have increased in feedlot cattle from 13.7% in 2011 to 17.8% in 2016 (Eastwood et al., 2017). It is estimated that liver abscesses can decrease carcass returns by $20 to $80 due to condemnation of the liver and nearby organs, depending on severity (Brown and Lawrence, 2010). An increase in liver abscesses has been linked to high concentrate grain diets during finishing (Nagaraja and Lechtenberg, 2007). The commonly accepted etiology is translocation of bacteria from an acidotic rumen (with rumen lesions as a predisposing factors) via the portal vein to the liver after a rumen wall perforation (Nagaraja and Lechtenberg, 2007). *Fusobacterium necrophorum* has classically been attributed as the causative organism of liver abscessation, though recently, the bacterial community within liver abscesses was found to be polymicrobial and diverse (Weinroth et al., 2017b). While feeding high grain diets is the primary predisposing factor to liver abscesses, other management factors, such as location of feedlot, location within feedlot, cattle type, length of time on feed, and other management strategies affect rates of abscessation.

In feedlot cattle, the main method of reducing liver abscesses is the use of antibiotics, namely tylosin phosphate or chlortetracycline. In 2011, tylosin was fed to an estimated 71% of all U.S. feedlot cattle housed in feedlots with a capacity of over 1000 animals (USDA–APHIS–VS–CEAH–NAHMS, 2013). When administered parentally, tylosin reduces prevalence and severity of liver abscesses (Nagaraja and Lechtenberg, 2007). Tylosin falls into a class of antibiotics known to inhibit bacterial protein synthesis (Nakajima, 1999); however, the exact

effect of tylosin on liver abscess rate reduction is unknown. It is hypothesized that tyolsin has an inhibitory effect on *F. necrophorum* or alters rumen bacteria which in turn reduces ruminal acidosis (Nagaraja and Lechtenberg, 2007).While tylosin does reduce prevalence and severity of liver abscesses, it is not consistent in reducing and eliminating abscesses altogether.

An emerging concern regarding tylosin use in feedlot cattle is selection for communities of bacteria that are more resistant to antibiotics. Both the Center for Disease Control and Prevention and the World Health Organization (**WHO**) have expressed concern regarding tylosin, with WHO classifying this class antimicrobials 'Highest Priority Critically Important Antimicrobial,' the most critical designation of an antimicrobial (World Health Organization, 2016). From there, the concern is that bacteria which harbor antibiotic resistant genetics (**AMR**) from foods of animal origins could infect humans (Center for Disease Control and Prevention, 2017b) and lead to treatment failure in the case of medical need. As a result of these concerns, the Food and Drug Administration implemented new regulation regarding administration of antimicrobials to food animals to reduce use of antimicrobials that are important in human healthcare.Macrolides, as one example, on January 1, 2017, were changed from over-the-counter to requiring a prescription as a result of the veterinary feed directive, which mandates that producers cannot use antimicrobials in feed and water off label (80 FR 31707). As a result, macrolides are still available for use, but require greater oversight in the valid veterinarian-client-patient-relationship.

The growing concern over the transmission of AMR to humans via food animals has resulted in an interest in replacing traditional husbandry practices with alternatives that do not require antimicrobials. Some of these alternatives include essential oils and probiotics, though none have been found to be as effective as an antimicrobial (Meyer et al., 2009; Pukrop et al.,

2017). In order to replace antimicrobials, the drug effects on both target and commensal bacteria must be understood. While there has been work on the effect of tylosin on the microbiome and resistome of beef cattle, the work was small in scope and performed on individual animals (Thomas et al., 2017). The objective of this study was to pair traditional culture techniques with metagenomic next generation sequencing to characterize microbiome and resistome differences among differing liver abscess rates from cattle administered or not administered tylosin in diverse U.S. feedlot locations.

Materials and Methods

Cattle Population and Experimental Design. Two thousand two hundred fifty-six cattle were enrolled in the study from 5 different feedlots in the United States encompassing BIFSCo regions 2 (comprised of California and Nevada) henceforth referred to as R2, 3 (comprised of Arizona, New Mexico and Texas) henceforth referred to as R3, and 4 (comprised of Colorado, Montana, Utah, and Wyoming) henceforth referred to as R4. Cattle were assigned to 16 pens with an average pen size of 141 head. The feedlot in R2 housed dairy-type cattle while the cattle in R3 and R4 were characterized as beef-type cattle, hence partially confounding region effects. The study was arranged in a randomized complete block design; with region used as the block (R2 and R3 both housing 4 pens of cattle and R4 housing 6 pens). In each block, an even number of pens were fed or not fed tylosin phosphate according to manufacture instructions (Elanco, Greenfield, IN). All cattle were fed and harvested during the same period and fed similar high concentrate grain diets besides the previously stated tylosin differences. Cattle base diets containing different combinations of sweet bran, cotton seed, fed straw, distillers grains, flaked corn, natural molasses. Cattle were treated for BVD and parasites at arrival, cattle in

conventional systems were fed Tylosin and Rumensin according to manufacture instructions throughout the feeding period.

Fecal and Soil Sample Collection. Both pen floor feces and soil samples were collected no more than 48 hours before harvest. Pen fecal samples were comprised of 20 (25g) pen floor fecal grabs collected along the outside and diagonals of each pen for a 500g composite per pen. Soil samples were collected in the same manner at 20 locations from the same pen floor and composited into one 500g sample. After collection, samples were placed on ice and transported to Colorado State University. Upon arrival, soil samples were frozen (-80°C) until further processing while a 25g aliquot of each fecal sample was removed from the composite samples before freezing (at -80°C) for culture work.

In Plant Data Collection. Within two days of fecal and soil collection, cattle were shipped to a commercial processing facility for harvest, where pen identification was maintained via tag transfer on the kill floor. Liver abscess incidence were recorded for all pens using the methods described by Brown and Lawrence (2010). Liver abscesses that exhibited an 'A-', 'A', or 'A+' liver score as defined by Eli Lilly (Elanco) Liver Check System were considered abscessed and recorded as such (severity was not recorded).

DNA Extraction. Ten grams of each soil and fecal sample were thawed and DNA was extracted using the Mo-Bio PowerMaxSoil DNA isolation kit (Mo Bio Laboratories, Inc., Solana Beach, CA) following manufacturer's protocols. Quality and concentration were evaluated using a NanoDrop™ spectrophotometer (Thermo Fisher Scientific, Inc.). Samples with a 260mm:280mm ratio equal to or higher than 1.8 and a concentration equal to or greater than 20 ng/µl of DNA were sequenced; samples not meeting these thresholds were concentrated using ethanol precipitation.

Shotgun Sequencing. One hundred µl aliquots of all soil and fecal DNA samples were delivered to the Genomic and Microarray Core at the University of Denver for library preparation and sequencing (Denver, CO). Sequencing libraries were constructed using an Illumina TruSeq DNA Library Kit (Illumina, Inc. San Diego, CA). Samples were sequenced (2x150) on four lanes of the Illumina HiSeq 4000 (Illumina, Inc. San Diego, CA); individual samples were divided across two lanes to alleviate lane effects.

16S Sequencing. Thirty microliter aliquots of DNA from fecal and soil samples were delivered to Novogene Bioinformatics Technology Company (Chula Vista, CA) for library preparation and sequencing. The V4 region of the 16S rRNA subunit was amplified with the 515F/806R primer set. Paired-end sequencing (2x250) was completed on two lanes of an Illumina HiSeq 2500 (Illumina, Inc. San Diego, CA).

Fecal Microbiological Analysis. All pen fecal samples were subjected to isolation of *Enterococcus* and generic *Escherichia coli*. The 25 g of retained feces from each pen was added to 225 ml of tryptic soy broth (**TSB**). Samples were incubated at 37°C for 24 h before plating onto Enterococcosel (**EC**) or MacConkey (**MAC**) agar. Enterococcosel plates were incubated for 48 h while MAC plates were incubated for 24 h, both at 43°C. Morphologically representative colonies from EC and MAC were re-streaked two times onto the same selective agars and incubated as described above. An isolate from each fecal sample from both EC and MAC was placed in TSB containing 10% glycerol and frozen (-80°C) until determination of antimicrobial susceptibility.

Minimum inhibitory concentrations (**MIC**) of isolated *Enterococcus* and generic *E. coli* to various antimicrobial drugs (Table 3.1) were determined using a broth microdilution method (Gragg et al., 2013). Minimum inhibitory concentrations and breakpoints were set using those

established by Clinical and Laboratory Standards Institute (2010) or the National Antimicrobial Resistance Monitoring System (2010).

16S Bioinformatics. Raw, demultiplexed reads, trimmed of sequencing adaptors, were received from Novagene. Using QIIME2 (Version 2-2018.2, Caporaso et al., 2010), reads were assigned exact sequence variants (**ESV**) via DADA2, using the paired end option with a 19 bp trimmed from the forward reads and 20bp from the reverse reads, both from the '5 end; the chimera detection method was set to 'pooled.' Fecal and soil samples were divided into two different datasets via 'feature-table filter-samples.' Fecal samples were rarified to a depth of 151,267 reads and soil samples to a depth of 141,106 reads. Taxonomic identity was determined using a pretrained Naïve Bayes classifier trained on the Greengenes database (v. 13_8, DeSantis et al., 2006) specific to the V4 region. Exact sequence variants assigned to mitochondria and chloroplasts, ESV's with "hits" to only one sample, and completely unknown taxa that did not have a Kingdom assignment were removed.

For construction and validation of the predictive model, 5 additional fecal and 5 soil samples and pen level liver abscesses percentages from the same BIFSCo regions were collected (during a different study), but a year later were included for validation of the prediction model. For construction of the predictive model, the five additional soil and fecal samples were assigned ESV and taxonomy using the same parameters and merged into the existing sample table via the 'merge' commands in the feature table plugin. The table (including both feces and soil hits for each sample) was rarified at 74,972 reads to allow for all samples to be included. Filtering to remove mitochondria, chloroplasts, and unassigned bacteria was performed. Sixteen of the samples were randomly chosen to serve as the training dataset while the remaining five were held out for model validation. The training model variable selection was performed using

LASSO via the 'coeff' function in the glmnet package (using the λ that gave the minimum mean cross-validated error via the 'cv.glmnet' function) and used to construct a linear model via 'lm'. The 'predict' function was used to estimate liver abscess percentages for the five samples held out from the training dataset. Root Mean Square Error (**RMSE**) was calculated with the 'rmse' function from the Metrics package.

Shotgun Bioinformatics. Demuliplexed fastq files from across different sampling lanes were concatenated together using the Linux 'cat' command so that each sample had one forward and one reverse fastq file. Resistome analysis was conducted via the AMRplusplus pipeline (Lakin et al., 2017). Briefly, samples were trimmed using Trimmomatic (Bolger et al., 2014a) and bovine DNA was removed from the trimmed reads using the Burrows Wheeler Aligner (**BWA**) (Li and Durbin, 2009) to align to the reference *Bos taurus* genome (UMD 3.1) as well as to the draft *Bos indicus* genome (Canavez et al., 2012). Any read that aligned to either genome were removed. Trimmed non-host DNA reads were aligned to the MEGARes' (Lakin et al., 2017) antimicrobial resistance database using BWA. Per Noyes et al. (2016), only AMR genes with gene fraction of >80% were considered present in a sample and included in further analyses. The number of hits to each AMR gene were compiled and each gene was assigned to an AMR class, mechanism, and group. For shotgun resistome analysis, genes present in less than 3 samples were removed. Remaining counts were normalized by cumulative sum scaling with a default percentile for normalization (Paulson et al., 2013b).

For a comparison of 16S rRNA amplicon sequencing results to shotgun results, the microbiome shotgun metagenomics results were obtained with the AMR++ pipeline (Lakin et al., 2017) using Kraken (Wood and Salzberg, 2014) for assigning taxonomic labels to reads.

Statistical analysis

Overall liver abscesses rates. When comparing cattle fed with tylosin versus those not, individual cattle (N = 2256) were considered instead the pen level differences as individual liver data was collected. Using PROC GENMOD SAS (version 9.4) with a Poisson distribution, liver abscess rate was compared between tylosin exposure groups controlling for feedlot differences and a correlated response of animals in the same pen. Alpha error was set at 0.05, while 'trends' were reported at between 0.051 and 0.10.

Shotgun statistics. To assess community-wide difference in resistomes between tylosin exposure or non-exposure groups, non-metric multidimensional scaling (NMDS) ordination was performed using Hellinger transformation and Euclidean distances using R (v. 3.3.0) with Vegan's 'metaMDS' (Legendre and Gallagher, 2001). These difference were formally compared using analysis of similarities (**ANOSIM**) (Clarke, 1993). Multivariate, zero-inflated Gaussian mixture models were fit to class and mechanism-level normalized counts using metagenomeSeq's 'fitZig'. (Paulson et al., 2013). Fitzig's output was used in limma's 'makeContrasts' and 'eBayes' functions to conduct pairwise comparisons of log2-fold change in abundance between tylosin treated and non-treated cattle, as well as between geographical regions, using $\alpha = 0.05$ with trends reported between 0.051 and 0.10. An inverse Simpson Diversity Index was used to evaluate richness between sample groups.

16S Statistics. For 16S rRNA amplicon sequencing microbiome analysis, alpha diversity (richness) was measured using Shannon's Entropy of Counts. Treatment alpha diversity was compared using a Kruskal-Wallis test via the 'qiime diversity alpha-group-significance.' Beta diversity differences were assessed using PERMANOVA (Anderson, 2001) with a weighted and unweighted unifrac distance matrix as the input (computed using "core-metrics'). Differences in

phyla and genus between treatment groups and regions were assessed using a Kruskal-Wallis test with a False Discovery Rate (**FDR**) correction via the 'group_significance.py' command in QIIME1.9 (Caporaso et al., 2010b). Alpha error was set at 0.05 and trends were reported between 0.051 and 0.10.

Comparing amplicon and shotgun metagenomic data. A Procrustes analysis was used to measure the correlation between the 16S rRNA amplicon and shotgun metagenomic sequencing results. Count tables were collapsed to the phylum level the 'metaMDS' function was applied to each matrix from the Vegan R package. The 'protest' function was used to test the non-randomness between the two configurations, with a visualization generated with the base 'plot' function.

Results

Overall Liver Abscess Rates and Tylosin Effect. Out of 2256 cattle, 199 liver abscesses were identified. Across all geographical regions, cattle that were not fed Tylosin were 2.1 (95% C.I. 1.5 to 2.8) times as likely to develop liver abscesses compared to those cattle that were fed tylosin. Cattle treated with tylosin had a lower ($P = 0.001$) liver abscess rate than those cattle not treated with tylosin. The reduction of liver abscesses via tylosin supplementation agreement with previous work (Nagaraja and Lechtenberg, 2007) and builds on the scientific consensus of this fact (Brown et al., 1975; Nagaraja and Chengappa, 1998).

Effect of Tylosin and Region on the Cattle Microbiome via Feces. On average, 16S rRNA amplicon fecal samples had 528,957 (Range 215,774 to 1,034,823) raw reads per sample. Seventy-two percent of reads were retained for down-steam analysis after ESV assignment via DADA2; resulting in an average of 382,666 assigned reads per sample (range 151,267 to

748,899). Across all samples, 33 phyla and 633 genera were identified, with an average of 22 phyla (range 16 to 29) and 277 genera (range 189 to 342) per sample.

Comparison of Treatment Groups and Region. Tylosin's impact on the microbiome of cattle was assessed via analysis of microbial communities in feces samples at the pen level. While liver abscess rate reduction via use of tylosin is well established (Nagaraja and Chengappa, 1998; Nagaraja and Lechtenberg, 2007) the impact of tylosin exposure on the fecal microbiome of cattle is less defined. Thomas et al. (2017) began the investigation of tylosin exposure, but their efforts were limited to a small sample size in one location. Here, tylosin exposure was compared across three different geographic regions and the sampling was comprised of over 2000 cattle; therefore, data from the present study provide a robust assessment of tyolsin and regional impact on the microbiome of feedlot cattle.

In the current study, alpha diversity did not differ between cattle that received tylosin versus those that did not ($P = 0.06$, Fig. 3.1B) or BIFSCo region of feeding ($P = 0.25$ Fig. 3.1A); however, numerically higher alpha diversity in cattle not fed tylosin was considered a trend ($P = 0.06$). Loss of bacterial community diverseness as a result of antibiotic treatment is well established, and has been exhibited both in human studies (Dethlefsen and Relman, 2011) and in the bovine rumen (Thomas et al., 2017). The lower alpha diversity of the fecal microbiomes of cattle treated with tylosin in this study further adds weight to supports past findings.

Fecal microbiome beta diversity, as measured by weighted and unweighted Unifrac distances, did not change between cattle treated and not treated with tylosin ($P = 0.16$, Fig. 3.2A and $P = 0.69$, Fig. 3.2B; respectively). However, weighted and unweighted Unifrac distances differed ($P = 0.001$, Fig. 3.2C and $P = 0.007$, Fig. 3.2D; respectively) between geographic regions. When regional differences were explored at the phyla and genus levels, there were no

differences ($P > 0.05$); this was also true for the comparison of tylosin fed and not fed cattle ($P > 0.05$) (Fig.3.3). The lack of beta diversity differences found in composite pen fecal microbiomes of tylosin treated and untreated cattle in or between individual phyla and genus was of interest because tylosin is known to act across a wide range of gram-positive organisms, but in the present study, there was no difference in gram-positive phyla. Not seeing a major change cattle associated microbiota between tylosin treated and non-treated cattle also has also previously been observed in rumen samples (Thomas et al., 2017).

Description of Regional Microbiome Differences via Soil Samples. On average, 16S rRNA amplicon sequenced soil samples contained 325,166 (Range 179,960 to 90,0612) raw reads per sample. Seventy-five percent of reads were retained for down-steam analysis after ESV assignment via DADA2, resulting in an average of 244,405 assigned reads per sample (range 141,106 to 667,975). Thirty-one phyla and 709 genera were detected, with an average of 23 phyla (range 17 to 29) and 283 genera (range 223 to 417).

Comparison of Treatment Groups and Region. Community composition of the soil from inside feedlot pens was the metric used to characterize regional microbiomes. Unsurprisingly, there was no difference in alpha diversity of soil samples between tylosin feed inclusion groups ($P = 0.44$, Fig. 3.4A), but BIFSCo regions of production differed in soil microbiome composition ($P = 0.01$, Fig. 3.4A). When soil microbiome composition by tylosin treatment group, when characterized using weighted or unweighted UniFrac distances, there was not a difference ($P = 0.13$ and $P = 0.24$, respectively), though geographical region beta diversity differed in both metrics ($P = 0.001$ and $P = 0.002$, respectively). When regional differences were further investigated at the phyla level, there were no differences ($P < 0.05$) in specific phyla, but Verrucomicrobia and Spirochaetes between regions were both tended ($P = 0.07$) to be different

between regions (Table 3.2). Verrucomicrobia have previously been described in bovine feces (Rice et al., 2012); and perhaps more interestingly, abundance has been reported to differ in feces by location of feeding (Weese and Jelinski, 2017). Spirochaetes also have been observed in bovine feces (Reti et al., 2013) and have been altered by different feeding strategies (Rice et al., 2012).These trends, in association with alpha diversity differences, likely drove differences in beta-diversity between geographical regions.

Microbiome differences between geographical production region, but not among differing antibiotic treatment groups, has been replicated in cow-calf herds in Western Canada (Weese and Jelinski, 2017). Weese and Jelinski (2017) found that differentiation of bacteria community membership was affected by location of farm, but not by antibiotics across farms—though when only one farm was considered, differences between treated and untreated calves were observed. These same patterns of regional difference, but without treatment differences, were observed in the present study; though sample size limited comparison of microbiomes specific to one farm.

Effect of Tylosin on the Antibiotic Resistance. Antibiotic Resistance was measured in two ways: phenotypically and genotypically. Genotype was described with shotgun metagenomics while phenotypic expression of resistance was assessed via expression of resistance genes in an indicator organism using MICs.

Culture Free. Shotgun metagenomics of composite fecal samples identified 12 classes, 13 mechanisms and 18 groups of antimicrobial resistance across all samples. Alpha diversity and overall resistome composition did not differ ($P = 0.19$ and $P = 0.46$, respectively) by treatment with or without tylosin (Fig. 3.5A). At the class level, there was no difference ($P > 0.05$) between relative abundance of AMR genes between treatment groups (Fig. 3.6). When evaluated

at the mechanism level, some mechanisms differed by treatment (namely Aminoglycoside N-acetyltransferases, Chloramphenicol acetyltransferases, and Undecaprenyl pyrophosphate phosphatase resistance). Of special note, macrolide resistance (of interest due to use of tylosin) at the mechanism level between treatments did not differ ($P = 0.16$) in relative abundance. Across all samples, genes aligning to three groups associated with macrolide resistance were detected: MphB, mefA, and msrD. The geographical BIFSCo region of production did result in differences ($P = 0.001$) in resistome composition (Fig 3.5B). Several classes of AMR in the feces of cattle in different geographical areas differed ($P > 0.05$, Table 3.3).

Tylosin exposure did not alter overall AMR richness or beta diversity or macrolide resistance. This finding agreed with Thomas et al. (2017), who found that there was no correlation between administration of antibiotics and AMR. This finding illustrates that exposure to an antibiotic does not always led to a linear increase in resistance when considered complex biological models, such as the feedlot pens investigated in this study. Other mechanisms of AMR not associated with any administered treatment differed between tylosin groups, such as Aminoglycoside N- acetyltransferases. Other studies also have reported aminoglycoside mechanisms to be present in cattle not administered that class of antibiotics, and changes to aminoglycoside resistance sans treatment with aminoglycosides (Noyes et al., 2016; Weinroth et al., 2018c). While not the objective of this study, this finding demonstrates the complexity of AMR genes in microbial ecology and that AMR is not solely driven by selection pressure.

Geographical BIFSCo region affected overall fecal resistome composition. Eight of ten classes of resistance genes that were detected differed by geographical region in pen composite feces resistome composition. While this finding was confounded by cattle source (in this case, cattle source was not standardized) differences did highlight that antimicrobial treatment is not

the only factor that contributes to AMR. It has been established that regional variation exists among resistomes in different regions of the world (World Health Organization, 2014).

Culture. No formal statistics were performed on the culture results; instead general trends were observed. In generic *E. coli*, isolates from the non-tylosin treated group did not display any intermediate or resistance to any of the 12 antimicrobials tested. The tylosin fed group did not demonstrate resistance or an intermediate breakpoint to 6 of the 12 antimicrobials tested: Amikacin, Cefoxitin, Chloramphenicol, Ciprofloxacin, Gentamicin, or Sulfamethoxazole-Trimethoprim. Of the remaining six antimicrobials, tylosin fed cattle had varying levels of resistance (Fig. 3.7C). *Enterococcus* isolates displayed resistance across both cattle fed and not fed treatments (Fig. 3.7A and 3.7B).

From a culture standpoint, *Enterococcus* isolates from feces of cattle fed tylosin were 100% (8/8) resistant to tylosin, while 25% (1/8 resistant and 1/8 susceptible) of isolates from feces from cattle that were not fed tylosin had an intermediate or resistant phenotype. These results agreed with those of Beukers et al. (2015), who found that resistance to tylosin increased in *Enterococcus* fecal isolates in cattle as a result of tylosin being provided in feed. Disparity between culture and shotgun results in the location of resistance genes has been described before (Weinroth et al., 2018c), and in that case (as well as likely in this study) was a result of extremely different methodology. While culture results are indicative of only one indicator organism, shotgun results allow for a survey of both culturable and unculturable bacterial genes; an ecological measure of resistance genetics prevalence. On the other hand, cultural methods verify phenotypic expression of a gene, as opposed to DNA sequencing which can only verify presence of the gene and not functional activity.

LASSO for Abscess Rate Prediction. Of all phyla identified across both soil and feces, six were selected for inclusion in the model by LASSO: four from feces (Euryarchaeota, Fibrobacteres, candidate phyla Cloacimonetes [WWE1], and WPS2) and two from the soil (Deferribacteres and Firmicutes). The training model was found to accurately predict the rate of occurrence (as a percentage of the total cattle in the pen) of liver abscesses (defined as an 'A-', 'A', or 'A+' liver score) within pens of feedlot cattle (adjusted $R^2 = 0.75$, residual standard error = 3.5, $P = 0.002$). When liver abscess rates from the five samples that were omitted from original analysis were used to predict liver abscess prevalence using the training model, the model accurately predicted the rate of occurrence of liver abscess in feedlot pens—albeit in only five samples (RMSE = 4.0, $R^2 = 0.88$, mean absolute error = 3.2).

The final model for estimating the rate of occurrence of liver abscesses within pens of feedlot cattle retained predictors of both fecal and soil phyla—indicating that animal feeding practices and geographical region of production both contribute to liver abscess prevalence in fed cattle. In composite pen floor feces samples, archaea Euryarchaeota and bacteria Fibrobacteres both have been previously described in bovine rumens (Jami and Mizrahi, 2012; Lange et al., 2005). Two candidate phyla in pen composite feces samples also were considered important in model: candidate phyla Cloacimonetes (which has been found in bovine feces [Sun et al., 2015]), and WPS-2 (found previously in ovine rumens (Wang et al., 2016). Euryarchaeota are known as methane producers in humans (Horz and Conrads, 2010), while Fibrobacteres and Cloacimonetes have both been associated with cellulose digestion (Limam et al., 2014; Ransom-Jones et al., 2012). Little is known about the function of candidate phyla WPS-2, though in addition to presence in ovine rumens, it has been found in anaerobic sludge digesters (Rivière et al., 2009) and soil (Hermans et al., 2016). The two soil associated phyla of interest included in

the model were Deferribacteres and Firmicutes. Deferribacteres, which are anaerobic respirators, were found in a wide variety of environments, but considered rare within many of these microbiomes (Alauzet and Jumas-Bilak, 2014). Firmicutes, on the other hand, have been widely reported at high concentrations in soil (Fierer et al., 2007) and bovine rumens (Jami et al., 2013; Jami and Mizrahi, 2012). Neither fecal nor soil Fusobacteria were selected for inclusion in the final model, which was unexpected as *Fusobacterium necrophorum* has typically been reported as the causative organism of liver abscessation (Nagaraja and Lechtenberg, 2007). It is important to note that, even with cross-validation during model construction and the implementation of a training and test dataset, this model was not based on use of a large sample size. While efforts were taken to reduce likelihood of overfitting, this model would benefit from an expanded sample size to bolster robustness of predictions.

Comparison of 16S rRNA Amplicon Sequencing to Shotgun Metagenomics. In this study, aliquots of the same DNA were prepared under when two different library preparation methods: 16S rRNA amplicon and shotgun metagenomics. Fecal shotgun metagenomic and 16 rRNA amplicon sequences libraries displayed different taxonomic relative abundance of phyla composition between the two library perpetration methods (Fig. 3.8A). Additionally, the number of phyla and orders across all samples found between the two preparations (Fig. 3.8B) was numerically higher in the 16S rRNA amplicon libraries. Procrustes analysis did not find a correlation ($P = 0.114$, $m^2 = 0.83$) between the two preparation methods at the phylum level. These findings are in agreement with Tessler et al. who also found 16S rRNA amplicon sequencing to find a greater taxonomic resolution of a bacterial community when compared to shotgun sequencing (Tessler et al., 2017).

Conclusions

Tylosin exposure was found to reduce the incidence of liver abscesses in feedlot cattle. Tylosin in feed did not alter cattle resistomes or microbiomes as measured by feces from this group. On the other hand, the geographical region cattle were raised in affected the cattle's fecal resistome. When pen floor soil samples were evaluated, unsurprisingly, geographical region had different soil resistomes but there were not differences between tylosin feeding groups in the soil. When soil and feces microbiome results were both used in a model to predict the incidence of liver abscesses in a pen of feedlot cattle, predictors from both soil and feces were included in the final model—indicating that cattle management and regional differences both drive the prevalence level of liver abscesses.

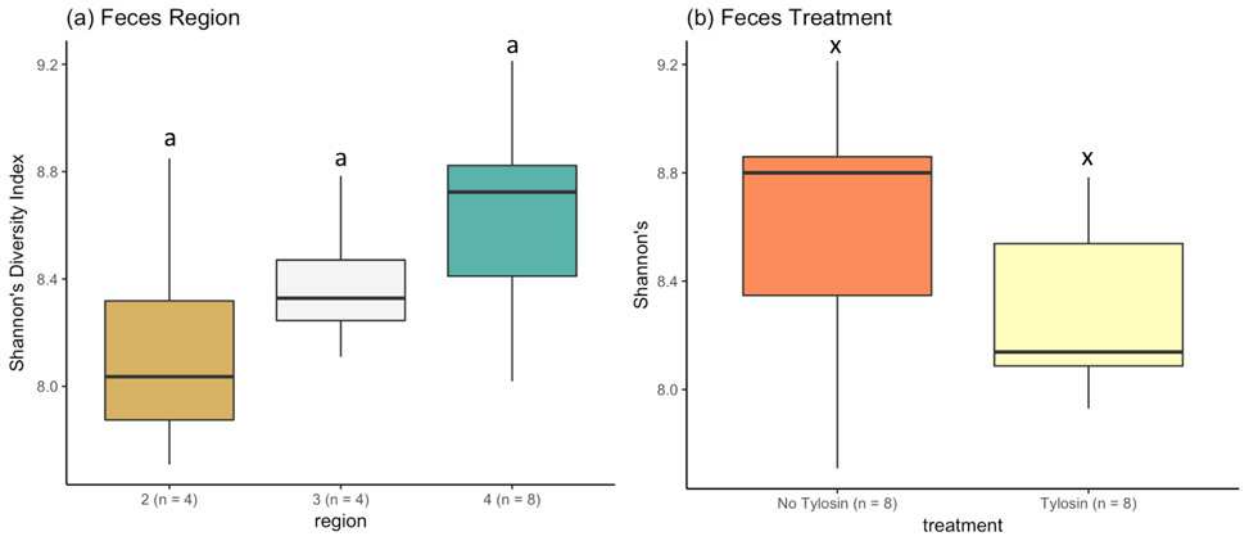


Figure 3.1: Cattle fecal microbiome alpha diversity measurements between (a) BIFSCo region and (b) cattle fed and not fed tylosin; while regions did not differ ($P = 0.25$) differences between tylosin feed and non-tylosin feed cattle was considered a trend ($P = 0.06$)
[abc, xy] Boxes that bear different superscripts within panel are different ($P < 0.05$).

Figure 3.2: weighted (a and c) and unweighted (b and d) Unifrac distances of the fecal microbiome of cattle colored by the feed inclusion of tylosin (a and b) and by BIFSCo region of feeding (c and d). While Tylosin feeding was not found to differentiate ($P < 0.05$) the microbiome composition, BIFSCo region of feeding did have an effect ($P > 0.05$).

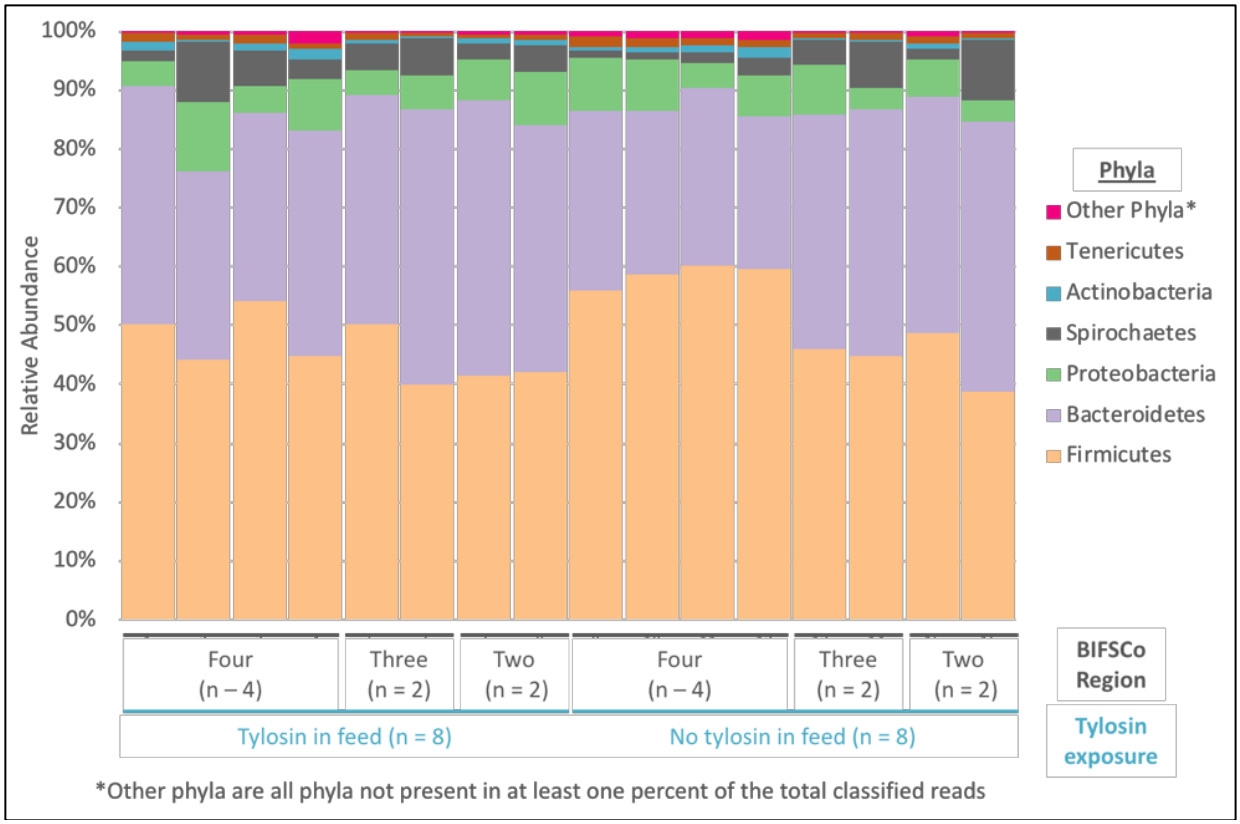


Figure 3.3: Relative abundance of the fecal microbiome by pen (organized by exposure to tylosin in feed and region of feeding). There were no differences ($P > 0.05$) between tylosin exposure group or treatment groups between individual phyla.

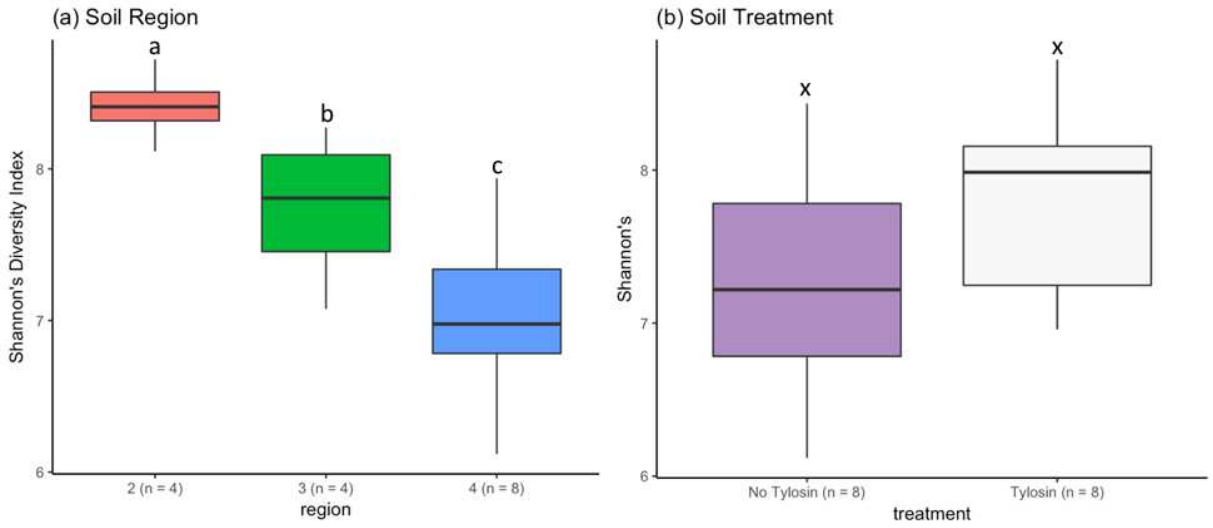


Figure 3.4: Soil alpha diversity measurements between (a) BIFSCo region and natural and (b) conventional cattle; while regions were different ($P = 0.01$), treatment with tylosin was not ($P = 0.44$).

$^{abc, xy}$ Boxes that bear different superscripts within panel are different ($P < 0.05$).

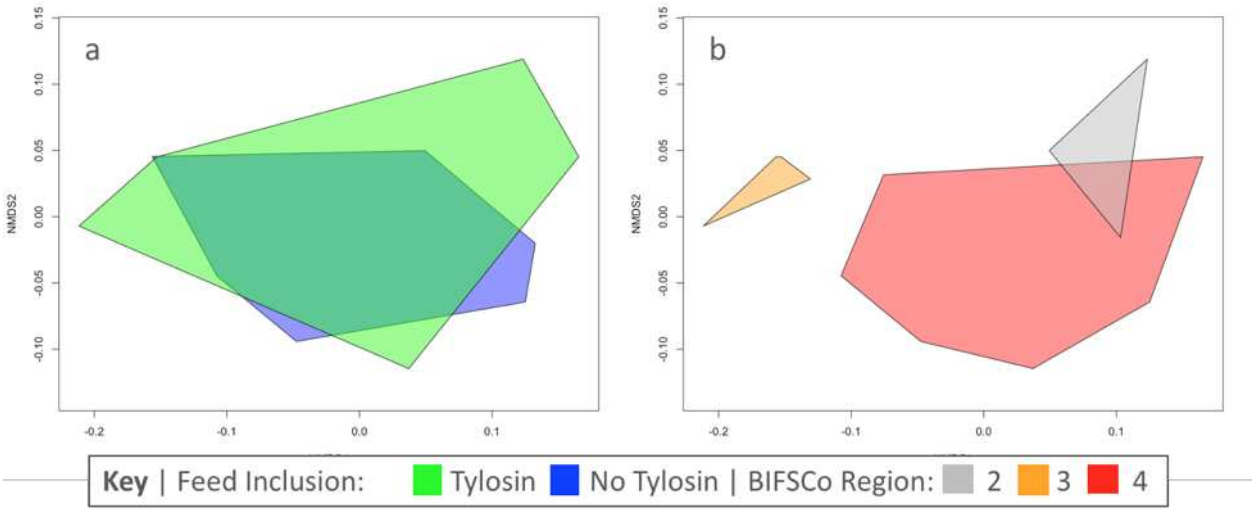


Figure 3.5: Non-metric multidimensional scaling (NMDS) ordination plot of resistome differences between (a) cattle supplemented with tylosin versus those not supplemented with tylosin (P = 0.46, Stress = 0.08, R= -0.01) and between BIFSCo regions of production (P = 0.001, Stress = 0.11, R= 0.48).

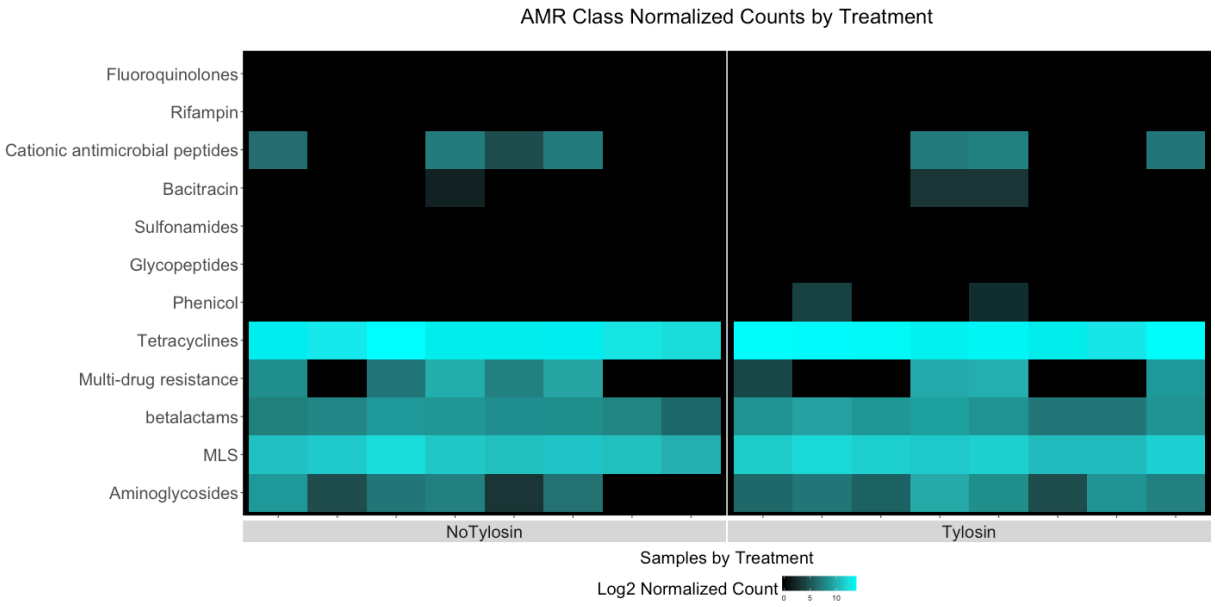


Figure 3.6: Heat map of Log2 normalized counts showing the abundance of classes of resistant antimicrobial genes found in all fecal pen samples. There was no difference ($P > 0.05$) between tylosin treated and untreated cattle overall or between classes.

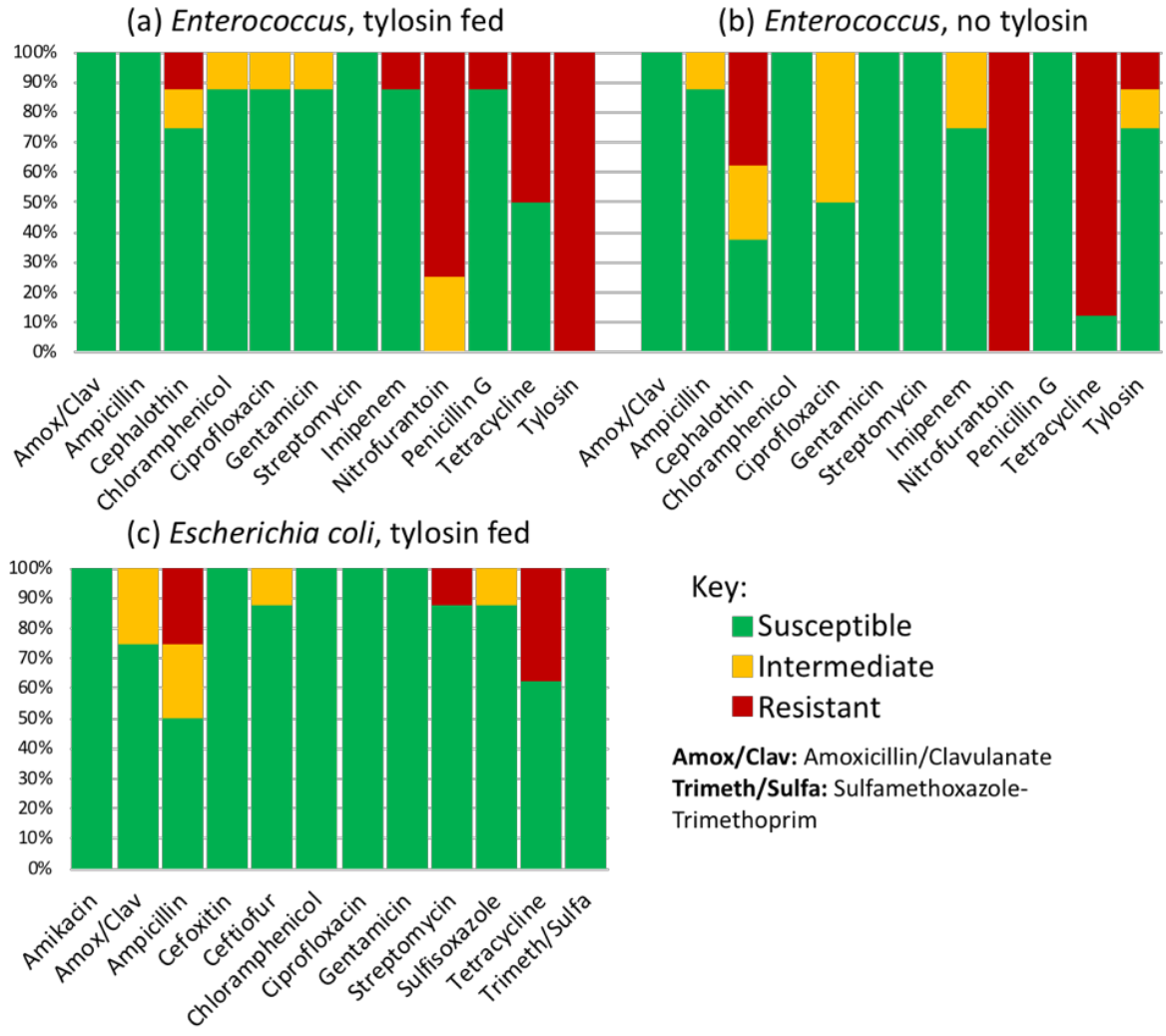
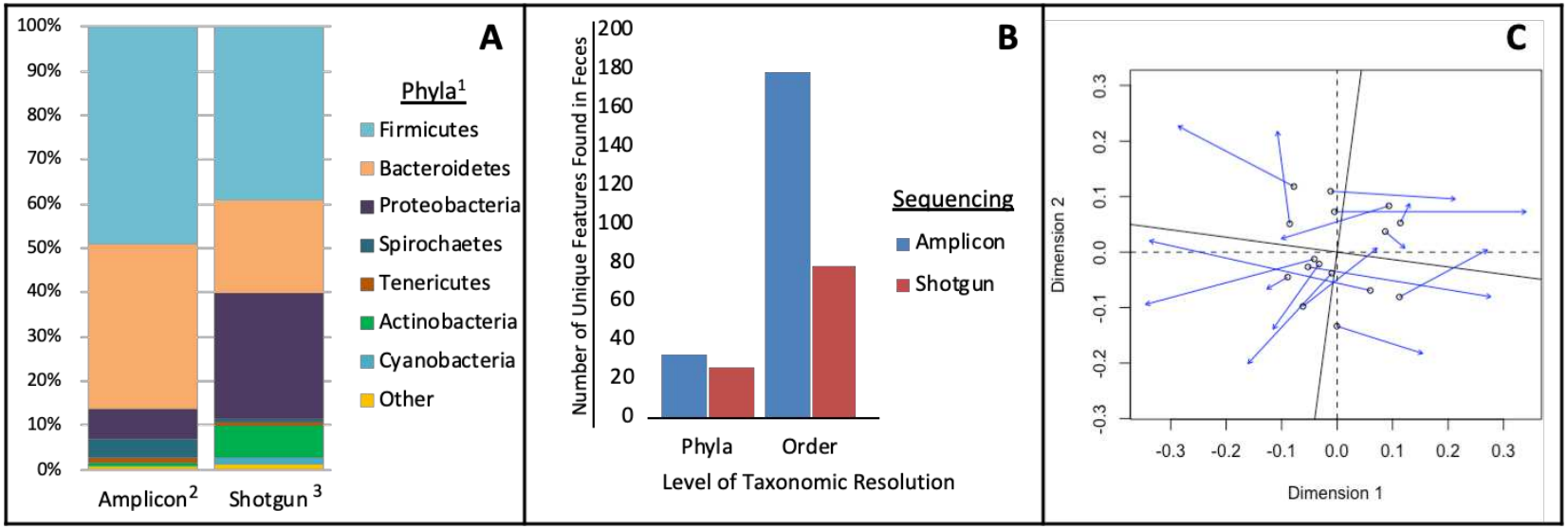


Figure 3.7: Susceptibility results of Enterococcus and generic Enterococcus cultured from pen floor cattle feces.



[1] Phyla identified in more than 1% of either 16 rRNA amplicon or shotgun metagenomic sequencing
[2] 16S amplicon rRNA Sequencing of the V4 Region
[3] Shotgun metagenomics

Figure 3.8: Comparison of two different library preparation methods (16S rRNA amplicon sequencing of the V4 region and shotgun metagenomics) of composite fecal samples by (A) taxonomic relative abundance of phyla present, (B) number of unique phyla and order of bacteria found across all fecal samples, and (C) Procrustes analysis of correlation ($P = 0.144$) between the two preparations.

Table 3.1: Antimicrobial drugs utilized for evaluating the susceptibility1 of Enterococcus, and generic *Escherichia coli* isolates.

Enterococcus	**Generic *E. coli***
Amoxicillin/Clavulanate	Amikacin
Ampicillin	Ampicillin
Cephalothin	Cefoxitin
Chloramphenicol	Ceftiofur
Ciprofloxacin	Chloramphenicol
Gentamicin	Ciprofloxacin
Streptomycin	Gentamicin
Imipenem	Streptomycin
Nitrofurantoin	Sulfamethoxazole-Trimethoprim
Penicillin G	Sulfisoxazole
Tetracycline	Tetracycline
Tylosin	

[1] Susceptibility was determined using breakpoints established by the Clinical and Laboratory Standards Institute (CLSI, 2013) or National Antimicrobial Resistance Monitoring System (NARMS, 2006).

Table 3.2: Number of rarified hits associated with Verrucomicrobia and Spirochaetes.

Phyla	FDR[1] P-Value	BIFSCo Region of Feedlot		
		Two	Three	Four
Verrucomicrobia	0.07	787	397	86
Spirochaetes	0.07	4892	3172	941

[1]False Discovery Rate

Table 3.3: Least squared means of normalized counts of classes of resistance by region of feeding, regardless of tylosin exposure

	BIFSCo Region of Feeding		
Class of Resistance	Two	Three	Four
Aminoglycosides	75[b]	350[a]	135[ab]
Bacitracin*	0[b]	5[a]	0[b]
Betalactams	401[a]	413[a]	211[b]
Cationic*	0[b]	150[a]	27[b]
Elfamycins	1365[a]	1576[a]	1444[a]
Fluoroquinolones*	0[b]	174[a]	87[ab]
MLS[1]	3378[a]	2438[b]	1830[b]
Multi-drug*	31[b]	853[a]	113[b]
Rifampin*	118[b]	257[ab]	290[a]
Tetracyclines	10949[a]	9335[a]	8254[a]

* Indicates one of more of the samples had zero hits attributed to that class of resistance.

[1] Macrolide-Lincosamide-Streptogramines.

[ab] Means that bear different superscripts within row are different ($P < 0.05$).

CHAPTER 4

16S rRNA Characterization of Liver Abscesses in Feedlot Cattle from Three States in the United States[1]

Summary

Liver abscesses are a major economic burden to beef producers. Although a few causative organisms have been cultured from purulent material, the full polymicrobial diversity of liver abscesses has not been reported. The objective of this study was to characterize purulent material collected from liver abscess in beef cattle produced in different production systems in three cattle producing states in the United States using 16S rRNA gene sequencing. Due to the data structure, a look into differences between region of feeding and application of a common antimicrobial was associated with microbiome composition was also conducted. Cattle included in the study were fed in California (dairy-type), Colorado and Texas (both beef-type). Liver abscesses from a cross-section of feedlots, geographic areas, and tylosin phosphate administered groups were collected at harvest, DNA from 34 liver abscess samples was extracted, and the V4 region of the 16S rRNA gene was amplified and sequenced. Sequences were classified into five phyla, 13 classes, and 17 orders in the domain Bacteria. The phyla identified included: Bacteroidetes (35.2% of reads), Proteobacteria (28.6%), Fusobacteria (18.2%), Firmicutes (12.4%) and Actinobacteria (5.5%). Sequences matching the genera *Fusobacterium* and *Trueperella*, which have previously been identified as causative agents in liver abscesses, were both present in the abscess bacterial communities at a rate of 15.1% and 3.2% of the overall

[1] This work has been previously published: M. D. Weinroth, C. R. Carlson, J. N. Martin, J. L. Metcalf, P. S. Morley, K. E. Belk, Journal of Animal Science, Volume 95, Issue 10, 1 October 2017, Pages 4520–4525, https://doi.org/10.2527/jas2017.1743

relative abundance, respectively. Furthermore, three of the most common phyla were gram-negative bacteria. An analysis of similarities (ANOSIM) test was conducted on Euclidean distances to assess differences between cattle treated and not treated with tylosin as well as to assess differences between regions. Geographical region and treatment with tylosin did affect the microbiome ($P = 0.002$ and $P = 0.026$ respectively), however a more robust sample scheme is needed to explore these differences. To our knowledge, this is the first publication describing the complex community of liver purulent material using next generation sequencing in cattle. These data may provide the framework for research on a more targeted approach to liver abscess prevention and treatment.

Introduction

The occurrence of a liver abscesses in feedlot cattle is associated with negative performance and economic impacts. It is estimated that liver abscesses, observed at prevalence of 20.9% of fed cattle harvested in the U.S., can decrease carcass value by $20 to $80 (Brown and Lawrence, 2010; McKeith et al., 2012). Classically, the primary etiology of this disease is attributed to *Fusobacterium necrophorum* (Nagaraja and Chengappa, 1998). However, there has been some ambiguity in compositional differences of other microflora's role in potential role in etiology that can be found in these abscesses, and these differences appear to vary systematically among feedlots and feeding strategies. Previous studies have indicated that liver abscesses are polymicrobial utilizing anaerobic and aerobic cultures (Nagaraja and Chengappa, 1998) and whole-genome-sequencing of isolated bacteria (Amachawadi et al., 2016). As reductions of liver abscesses remain a concern for the industry, and the use of antimicrobial drugs used for prevention and treatment are increasingly scrutinized, a more thorough understanding of the bacteriology of liver abscessation is warranted. Therefore, this study characterizes the microbial

communities in the purulent materials of liver abscesses using 16S rRNA gene amplicon sequencing.

Materials and Methods

Cattle Population. Sixteen pens of feedlot cattle (Average number of animal in pen: 141), from 5 different feedlots in the United States (one feedlot in California, one feedlot in Texas, and three feedlots in Colorado) were utilized for this study. Cattle included in the study were a mix of dairy-type (California) and beef-type (Colorado and Texas). In order to investigate the effects of a commonly utilized liver abscess control strategy, one half of the enrolled pens (n = 8) housed cattle that were supplemented with tylosin phosphate following label usage (Elanco, Greenfield, IN) for the duration of the feeding period, while the other eight pens were not fed tylsoin. The pens were identified prior to slaughter to facilitate sample collection at the time of harvest.

Liver abscess collection. A sample of livers identified as having abscess (up to 5 per pen) were reserved for removal of the abscess. Liver abscess collection was performed by removing the abscess from the liver using a sterile scalpel; taking care to avoid puncturing the abscess during the collection process. When multiple liver abscesses were present in one liver, the abscess collected was the most convenient to collect that appeared to harbor the most purulent material. Liver abscess samples were placed in sterile bags (Whirl-Pak; Nasco Corp., Fort Atkinson, WI) and transported on ice to Colorado State University. An aliquot of the abscess purulent material was sterily removed from the abscess, placed in sterile 50 ml conical tubes (Thermo Fisher Scientific, Waltham, MA), and frozen (-80°C) until the time of DNA Extraction. While the original sample plan called for the collection of 80 liver abscess, due to logistical limitations (some pens of cattle did not have 5 abscessed livers to sample) and the limit on the

amount of purulent material from some abscesses (resulting from the stage and size of the abscess), DNA from 34 liver samples was successful extracted and used in downstream analysis.

DNA Extraction and Sequencing. DNA was extracted from 0.021 g to 0.725 g of purulent material using the Mo-Bio PowerFecal DNA isolation kit (Mo Bio Laboratories, Inc., Solana Beach, CA) following manufacturer's protocols. Quality and concentration was evaluated using a NanoDropTM spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA). DNA extraction samples with a 260mm:280mm ratio < 1.3 (lab average ratio 1.80) and a concentration < 20 ng/µl of DNA (concentration average 21.8 ng/µl) were concentrated using ethanol precipitation prior to sequencing.

Thirty microliter aliquots of DNA from all liver samples were delivered to Novogene Bioinformatics Technology Company (Chula Vista, CA) for library preparation and sequencing. The V4 region of the 16S rDNA subunit was amplified with the 515F/806R primer set. Paired-end sequencing (2 x 250 base pairs) was completed on the Illumina HiSeq 2500 (Illumina, Inc. San Diego, CA).

Bioinformatics and Statistical Analysis. Reads from sequencing adaptors were trimmed from raw sequence data using cutadapt (Martin, 2011). Forward and reverse reads for each sample were merged using PEAR v0.9.10 (Zhang et al., 2014) with a minimum read length 187 and a maximum of 310. Using QIIME v1.9 (Caporaso et al., 2010b), raw sequencing reads were clustered into OTU (Operational Taxonomic Units) using open reference methods at 97% similarity. De novo OTU were clustered using sumaclust (Mercier et al., 2013)while reference-based clustering relied on sortmerna (Kopylova et al., 2012) and the Greengenes 16S rRNA reference database (DeSantis et al., 2006a).Taxonomy was assigned to OTU with UCLUST (Edgar, 2010) using the Greengene database. Operational taxonomic units assigned to

mitochondria and chloroplasts and singleton OTU were removed. The OTU table was normalized using cumulative sum scaling. A rarefaction curve was constructed with using Chao1 measurements from biom files to assess whether sequencing depth sufficiently captured diversity. Taxa present in all samples a relative abundance of at least 1.0% in all samples were considered part of the 'common microbiota.'

Non-metric multidimensional scaling (NMDS) ordination using Euclidean distances were calculated on cumulative sum scaled normalized counts and the analysis of similarities "ANOSIM" function in the Vegan Package in R was used to assess differences between cattle treated and not treated with tylosin as well as to assess differences between regions. For all comparisons in the study, $\alpha = 0.05$ was utilized.

Results and Discussion

Rarefaction Curve. The total number of reads considered in the analysis after quality control was 13,505,094. The mean number of reads in each sample was 350,716 (range: 190,551 to 996,910; standard deviation 191,995). The rarefaction curve (Fig 4.1) shows a plateauing of reads mapped to novel OTU as the proportion of reads sampled increases, indicating that the microbial community was sampled to an appropriate depth to allow complete characterization of the bacterial community.

Summary Statistics. Through 16S characterization, 5 phyla, 13 classes, and 17 orders were identified in the DNA extracted among all sampled abscesses (Fig. 4.2). The phyla identified were: Bacteroidetes (35.2% of reads), Proteobacteria (28.6%), Fusobacteria (18.2%), Firmicutes (12.4%) and Actinobacteria (5.5%). Of note, reads mapped to the three most predominant phyla, which comprised 82% of the abscess microbial community, represent gram-negative bacteria. This result is interesting as tylosin, a macrolide commonly used to prevent

liver abscesses in feedlot cattle, is primarily active against gram-positive bacteria with limited efficacy against gram-negatives. While the isolation of gram-negative bacteria from liver abscesses has been described (Nagaraja and Chengappa, 1998), the abundance of gram-negative bacteria in the gut microbial community may explain why common prevention strategies (i.e. macrolide supplementation) have imperfect efficacy for prevention of liver abscessation. Instead, these results suggest that reduction strategies which provide more broad spectrum action against the entire microbiome may be beneficial in more completely preventing this disease.

Identification of bacteria of interest and common microbiota. Previous literature has identified *Fusobacterium necrophorum* as the primary causative organism for bovine liver abscesses (Nagaraja and Chengappa, 1998). *Fusobacterium* was found in all liver abscesses sampled in this study, and on average reads mapped to this genus comprised 15.1% of the microbial community when characterized using 16S rRNA amplicon sequencing (range: 10.6% to 21.9%). Another common bacteria associated with liver abscesses, *Trueperella pyogenes*, was identified at genus level in all of the samples but at a lower overall community makeup of 3.2% (range: 2.4% to 5.1%).

Ten other bacterial genera were present in all samples, 5 at a greater relative abundance than 3% of all classified reads: *Bacteroides* (17.6% of mapped reads), *Porphyromonas* (14.1%), *Pseudomonas (*5.7%), *Enterobacteriaceae* (3.7%; classified at the family level), and *Sneathia* (3.1%). The remaining five genera were present in 2.2 to 2.9% of the relative abundance of the community: *Parvimonas* (2.9%), *Helcococcus* (2.8%), *Psychrobacter* (2.6%), *Atopobium* (2.4%), *Campylobacter* (2.2%), and *Haemophilus* (2.2%). This bacterial community shares several genera (namely *Bacteroides*, *Enterobacteriaceae*, and *Fusobacterium)* with a characterization of human liver abscesses also done with 16S rRNA sequencing, though other genera, such as

Klebsilla, were highly represented in the human abscesses and not characterized in this analysis (Song et al., 2014).

Bacteroides and *Porphyromonas* have both been previously described as being present in bovine liver abscess purulent material (Nagaraja and Lechtenberg, 2007; Scanlan and Hathcock, 1983). While *Pseudomonas* has been reported in other ruminant abscesses (Tadayon et al., 1980) and *Sneathia* has been linked to abscesses in the cervical lymph nodes of other mammals (Eisenberg et al., 2016). Organisms from the Enterobacteriaceae family have also been cultured from liver purulent material such as *Salmonella enterica* (Amachawadi and Nagaraja, 2015).

While there are no previous reports regarding the presence of *Parvimonas*, *Helcococcus, Psychrobacter*, *Atopobium, Campylobacter* or *Haemophilus* in peer-reviewed literature as a causative agent of bovine liver abscesses, several of these species have been associated with other related disease. For examples, co-occurrence of *Parvimonas* and *Fusobacterium* has been found in human colorectal cancer (Nakatsu et al., 2015), *Helcococcus ovis* has been reported in associated with bovine valvular endocarditis (Kutzer et al., 2008),and *Atopobium* has been found in purulent material of other mammals (Oyaert et al., 2014).

Initial Evaluation of Differences Related to Geographic Region and Tylosin Exposure. Though not the primary objective of the study, and limited by sample size, sampling structure used in this study provided an opportunity for initial investigation of differences that might exist in the flora of liver abscesses based upon comparison of geographic regions (n = 8 California, n = 9 Texas, and n = 17 Colorado) and tylosin exposure (n = 18 non-tylosin fed, n = 16 tylosin fed). Sampling size and the partial confounding of region by cattle type (all the cattle in California were dairy-type while the cattle in the other two regions were beef-type) limit extensive formal comparison. However, in an observation of feedlot location and

supplementation, region where feedlots were located, and supplementation with tylosin phosphate both affected liver abscess communities ($P = 0.002$ and $P = 0.026$, respectively). These comparisons among observed groups of cattle raise the possibility of different liver abscess rate composition by area and rearing methods; which in turn may lead to more targeted approaches to reduce abscess rates.

Conclusions

Management of liver abscesses in feedlot cattle continues to be an important priority for the North American beef cattle industry. Currently, the most common management strategies, which utilize treatment of cattle by including antimicrobial drugs in feed, are facing intensifying scrutiny. As such, a more thorough understanding of the microbial drivers of liver abscessation may lead to more efficient and sustainable management strategies; for example, the relationship between epimural bacteria and liver purulent material. The observed differences in region and rearing strategy provide further avenues for investigation and prevention of abscesses. We believe this characterization will allow for an ecological approach to treatment and prevention of liver abscesses.

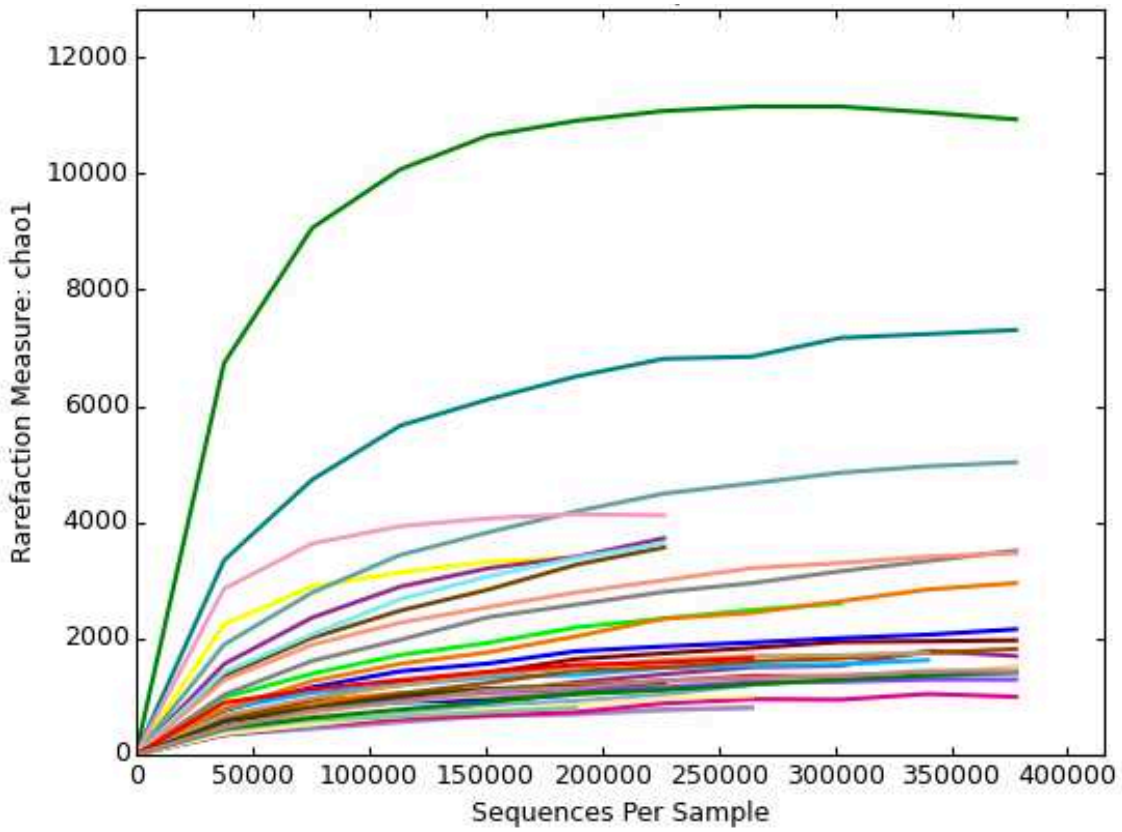


Figure 4.1: Rarefaction plot of each individual liver sample from chao1 measurements. The leveling off in all samples indicates an appropriate sampling depth was researched to estimate the diversity of the community.

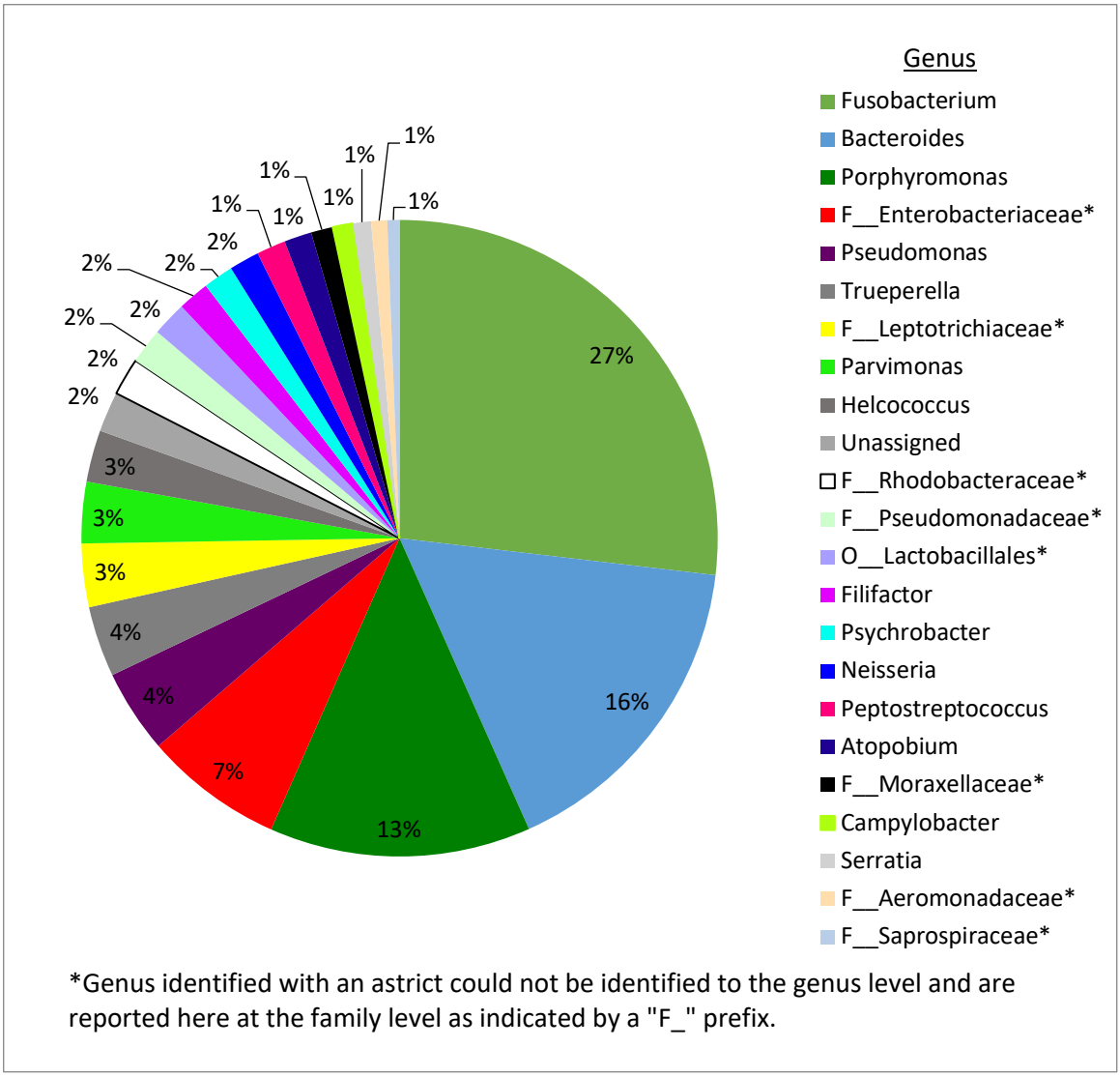


Figure 4.2: Genus level breakdown of liver abscess purulent material averaged over 34 liver purulent material samples taken in the study and sequenced with the V4 region of the 16S rRNA gene.

REFERENCES

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 44, W3–W10. doi:10.1093/nar/gkw343.

Agga, G. E., Schmidt, J. W., and Arthur, T. M. (2016). Antimicrobial-Resistant Fecal Bacteria from Ceftiofur-Treated and Nonantimicrobial-Treated Comingled Beef Cows at a Cow–Calf Operation. *Microb. Drug Resist.* 22, 598–608. doi:10.1089/mdr.2015.0259.

Aird, D., Ross, M. G., Chen, W.-S., Danielsson, M., Fennell, T., Russ, C., et al. (2011). Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* 12, R18. doi:10.1186/gb-2011-12-2-r18.

Alauzet, C., and Jumas-Bilak, E. (2014). "The Phylum Deferribacteres and the Genus Caldithrix" in *The Prokaryotes* (Springer, Berlin, Heidelberg), 595–611. doi:10.1007/978-3-642-38954-2_162.

Althani, A., Marei, H., Hamdi, W., Nasrallah, G., El Zowalaty, M., Khodor, S., et al. (2016). Human Microbiome and its Association With Health and Diseases. *J. Cell. Physiol.* 231, 1688–1694.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi:10.1016/S0022-2836(05)80360-2.

Amachawadi, R. G., and Nagaraja, T. G. (2015). First Report of Anaerobic Isolation of Salmonella enterica from Liver Abscesses of Feedlot Cattle. *J. Clin. Microbiol.* 53, 3100–3101. doi:10.1128/JCM.01111-15.

Amachawadi, R. G., Thomas, M., Nagaraja, T. G., and Scaria, J. (2016). Genome Sequences of Salmonella enterica subsp. enterica Serovar Lubbock Strains Isolated from Liver Abscesses of Feedlot Cattle. *Genome Announc.* 4. doi:10.1128/genomeA.00319-16.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* 26, 26–46. doi:https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x.

Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* 6, 343–351. doi:10.1038/ismej.2011.119.

Bayat, A. (2002). Bioinformatics. *BMJ* 324, 1018–1022.

Beaber, J. W., Hochhut, B., and Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427, 72–74. doi:10.1038/nature02241.

Benjamin, Y., and Hochberg, Y. (1995). Controlling for the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc.* 57, 289–300.

Berendonk, T. U., Manaia, C. M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., et al. (2015). Tackling antibiotic resistance: the environmental framework. *Nat. Rev. Microbiol.* 13, 310–317. doi:10.1038/nrmicro3439.

Bergholz, T. M., Moreno Switt, A. I., and Wiedmann, M. (2014). Omics approaches in food safety: fulfilling the promise? *Trends Microbiol.* 22, 275–281. doi:10.1016/j.tim.2014.01.006.

Beukers, A. G., Zaheer, R., Cook, S. R., Stanford, K., Chaves, A. V., Ward, M. P., et al. (2015). Effect of in-feed administration and withdrawal of tylosin phosphate on antibiotic resistance in enterococci isolated from feedlot steers. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00483.

Boeckel, T. P. V., Brower, C., Gilbert, M., Grenfell, B. T., Levin, S. A., Robinson, T. P., et al. (2015). Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci.* 112, 5649–5654. doi:10.1073/pnas.1503141112.

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

Bolger, A. M., Lohse, M., and Usadel, B. (2014b). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi:10.1093/bioinformatics/btu170.

Brichta-Harhay, D. M., Arthur, T. M., Bosilevac, J. M., Kalchayanand, N., Shackelford, S. D., Wheeler, T. L., et al. (2011). Diversity of Multidrug-Resistant Salmonella enterica Strains Associated with Cattle at Harvest in the United States. *Appl Env. Microbiol* 77, 1783–1796. doi:10.1128/AEM.01885-10.

Brito, I. L., and Alm, E. J. (2016). Tracking Strains in the Microbiome: Insights from Metagenomics and Models. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00712.

Brown, H., Bing, R. F., Grueter, H. P., McAskill, J. W., Cooley, C. O., and Rathmacher, R. P. (1975). Tylosin and chlortetracycline for the prevention of liver abscesses, improved weight gains and feed efficiency in feedlot cattle. *J. Anim. Sci.* 40, 207–213.

Brown, M. H., Gill, C. O., Hollingsworth, J., Nickelson, R., Seward, S., Sheridan, J. J., et al. (2000). The role of microbiological testing in systems for assuring the safety of beef. *Int. J. Food Microbiol.* 62, 7–16. doi:10.1016/S0168-1605(00)00408-6.

Brown, T. R., and Lawrence, T. E. (2010). Association of liver abnormalities with carcass grading performance and value. *J. Anim. Sci.* 88, 4037–4043. doi:10.2527/jas.2010-3219.

Brul, S., Bassett, J., Cook, P., Kathariou, S., McClure, P., Jasti, P. R., et al. (2012). 'Omics' technologies in quantitative microbial risk assessment. *Trends Food Sci. Technol.* 27, 12–24. doi:https://doi.org/10.1016/j.tifs.2012.04.004.

Bulut, E., Arthur, T. M., Schmidt, J. M., and Wang, B. (2017). Tracking Pre-Harvest Contamination Sources of Salmonella in Beef Products on a Quantitative Basis. in *Research Updates* (Houston, Texas: National Cattlemen's Beef Association).

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi:10.1038/nmeth.3869.

Canavez, F. C., Luche, D. D., Stothard, P., Leite, K. R. M., Sousa-Canavez, J. M., Plastow, G., et al. (2012). Genome Sequence and Assembly of Bos indicus. *J. Hered.* 103, 342–348. doi:10.1093/jhered/esr153.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010a). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi:10.1038/nmeth.f.303.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010b). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi:10.1038/nmeth.f.303.

Center for Disease Control and Prevention (2013). ANTIBIOTIC RESISTANCE THREATS in the United States, 2013. US Department of Health and Human Services Available at: http://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf.

Center for Disease Control and Prevention (2017a). Antibiotic Resistance, Food, and Food-Producing Animals. *Cent. Dis. Control Prev.* Available at: https://www.cdc.gov/features/antibiotic-resistance-food/index.html [Accessed December 26, 2017].

Center for Disease Control and Prevention (2017b). Antibiotic Resistance, Food, and Food-Producing Animals. *Cent. Dis. Control Prev.* Available at: https://www.cdc.gov/features/antibiotic-resistance-food/index.html [Accessed December 10, 2017].

Chambers, L., Yang, Y., Littier, H., Ray, P., Zhang, T., Pruden, A., et al. (2015). Metagenomic Analysis of Antibiotic Resistance Genes in Dairy Cow Feces following Therapeutic Administration of Third Generation Cephalosporin. *PLOS ONE* 10, e0133764. doi:10.1371/journal.pone.0133764.

Chouvarine, P., Wiehlmann, L., Moran Losada, P., DeLuca, D. S., and Tümmler, B. (2016). Filtration and Normalization of Sequencing Read Data in Whole-Metagenome Shotgun Samples. *PLoS ONE* 11. doi:10.1371/journal.pone.0165015.

Clarke, K. (1993). Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* 18, 117-`43.

Clinical and Laboratory Standards Institute (2010). *Performance standards for antimicrobial susceptibility testing; Twentieth Informational Supplement*. Wayne, PA: Clinical and Laboratory Standards Institute.

Crofts, T. S., Gasparrini, A. J., and Dantas, G. (2017). Next-generation approaches to understand and combat the antibiotic resistome. *Nat. Rev. Microbiol.* 15, 422–434. doi:10.1038/nrmicro.2017.28.

Culligan, E. P., and Sleator, R. D. (2016). Editorial: From Genes to Species: Novel Insights from Metagenomics. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.01181.

D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., et al. (2011). Antibiotic resistance is ancient. *Nature* 477, 457–461. doi:10.1038/nature10388.

Deng, X., Bakker, H. C. den, and Hendriksen, R. S. (2016). Genomic Epidemiology: Whole-Genome-Sequencing–Powered Surveillance and Outbreak Investigation of Foodborne Bacterial Pathogens. *Annu. Rev. Food Sci. Technol.* 7, 353–374. doi:10.1146/annurev-food-041715-033259.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006a). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5073.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006b). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072. doi:10.1128/AEM.03006-05.

Dethlefsen, L., and Relman, D. A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc. Natl. Acad. Sci.* 108, 4554–4561. doi:10.1073/pnas.1000087107.

Djordjevic, S. P., Stokes, H. W., and Chowdhury, P. R. (2013). Mobile elements, zoonotic pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans, production animals and soil microbiota. *Front. Microbiol.* 4. doi:10.3389/fmicb.2013.00086.

Doyle, C. J., O'Toole, P. W., and Cotter, P. D. (2017). Metagenome-based surveillance and diagnostic approaches to studying the microbial ecology of food production and processing environments. *Environ. Microbiol.* Preprint. doi:10.1111/1462-2920.13859.

Eastwood, L. C., Boykin, C. A., Harris, M. K., Arnold, A. N., Hale, D. S., Kerth, C. R., et al. (2017). National Beef Quality Audit-2016: Transportation, mobility, and harvest-floor assessments of targeted characteristics that affect quality and value of cattle, carcasses, and by-products1. *Transl. Anim. Sci.* 1, 229–238. doi:10.2527/tas2017.0029.

Economou, V., and Gousia, P. (2015). Agriculture and food animals as a source of antimicrobial-resistant bacteria. *Infect. Drug Resist.* 8, 49–61. doi:10.2147/IDR.S55778.

Eisenberg, T., Glaeser, S. P., Ewers, C., Semmler, T., Drescher, B., and Kämpfer, P. (2016). Caviibacter abscessus gen. nov., sp. nov., a member of the family Leptotrichiaceae isolated from guinea pigs (Cavia porcellus). *Int. J. Syst. Evol. Microbiol.* 66, 1652–1659. doi:10.1099/ijsem.0.000922.

Ercolini, D. (2004). PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. *J. Microbiol. Methods* 56, 297–314. doi:10.1016/j.mimet.2003.11.006.

Ercolini, D. (2013). High-Throughput Sequencing and Metagenomics: Moving Forward in the Culture-Independent Analysis of Food Microbial Ecology. *Appl. Environ. Microbiol.* 79, 3148–3155. doi:10.1128/AEM.00256-13.

Ercolini, D., Ferrocino, I., Nasi, A., Ndagijimana, M., Vernocchi, P., Storia, A. L., et al. (2011). Monitoring of Microbial Metabolites and Bacterial Diversity in Beef Stored under Different Packaging Conditions. *Appl. Environ. Microbiol.* 77, 7372–7381. doi:10.1128/AEM.05521-11.

Exploring the Sources of Bacterial Spoilers in Beefsteaks by Culture-Independent High-Throughput Sequencing Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0070222#pone-0070222-g002 [Accessed September 2, 2017].

Fierer, N., Bradford, M. A., and Jackson, R. B. (2007). Toward an Ecological Classification of Soil Bacteria. *Ecology* 88, 1354–1364. doi:10.1890/05-1839.

Finn, R. D., Clements, J., Arndt, W., Miller, B. L., Wheeler, T. J., Schreiber, F., et al. (2015). HMMER web server: 2015 update. *Nucleic Acids Res.* 43, W30–W38. doi:10.1093/nar/gkv397.

Forsberg, K. J., Patel, S., Gibson, M. K., Lauber, C. L., Knight, R., Fierer, N., et al. (2014). Bacterial phylogeny structures soil resistomes across habitats. *Nature* 509, 612–616. doi:10.1038/nature13377.

Galia, W., Leriche, F., Cruveiller, S., Garnier, C., Navratil, V., Dubost, A., et al. (2017). Strand-specific transcriptomes of Enterohemorrhagic Escherichia coli in response to interactions with ground beef microbiota: interactions between microorganisms in raw meat. *BMC Genomics* 18. doi:10.1186/s12864-017-3957-2.

Giaouris, E., Heir, E., Desvaux, M., Hébraud, M., Møretrø, T., Langsrud, S., et al. (2015). Intra- and inter-species interactions within biofilms of important foodborne bacterial pathogens. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.00841.

Goecks, J., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* 11, R86. doi:10.1186/gb-2010-11-8-r86.

Gragg, S. E., Loneragan, G. H., Brashears, M. M., Arthur, T. M., Bosilevac, J. M., Kalchayanand, N., et al. (2013). Cross-sectional Study Examining Salmonella enterica

Carriage in Subiliac Lymph Nodes of Cull and Feedlot Cattle at Harvest. *Foodborne Pathog. Dis.* 10, 368–374. doi:10.1089/fpd.2012.1275.

Hall, R. M., and Schwarz, S. (2016). Resistance gene naming and numbering: is it a new gene or not? *J. Antimicrob. Chemother.* 71, 569–571. doi:10.1093/jac/dkv351.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, R245–R249. doi:10.1016/S1074-5521(98)90108-9.

Hardwick, S. A., Chen, W. Y., Wong, T., Kanakamedala, B. S., Deveson, I. W., Ongley, S. E., et al. (2018). Synthetic microbe communities provide internal reference standards for metagenome sequencing and analysis. *Nat. Commun.* 9, 3096. doi:10.1038/s41467-018-05555-0.

Harwood, V. J., Levine, A. D., Scott, T. M., Chivukula, V., Lukasik, J., Farrah, S. R., et al. (2005). Validity of the Indicator Organism Paradigm for Pathogen Reduction in Reclaimed Water and Public Health Protection. *Appl. Environ. Microbiol.* 71, 3163–3170. doi:10.1128/AEM.71.6.3163-3170.2005.

Heather, J. M., and Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics* 107, 1–8. doi:10.1016/j.ygeno.2015.11.003.

Hermans, S. M., Buckley, H. L., Case, B. S., Curran-Cournane, F., Taylor, M., and Lear, G. (2016). Bacteria as emerging indicators of soil condition. *Appl Env. Microbiol*, AEM.02826-16. doi:10.1128/AEM.02826-16.

Hobbs, E. T., Pereira, T., O'Neill, P. K., and Erill, I. (2016). A Bayesian inference method for the analysis of transcriptional regulatory networks in metagenomic data. *Algorithms Mol. Biol. AMB* 11. doi:10.1186/s13015-016-0082-8.

Hoorfar, J. (2011). Rapid detection, characterization, and enumeration of foodborne pathogens. *APMIS* 119, 1–24. doi:10.1111/j.1600-0463.2011.02767.x.

Horz, H.-P., and Conrads, G. (2010). The Discussion Goes on: What Is the Role of Euryarchaeota in Humans? *Archaea*. doi:10.1155/2010/967271.

Hsu, P. D., Lander, E. S., and Zhang, F. (2014). Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 157, 1262–1278. doi:10.1016/j.cell.2014.05.010.

Hu, Y., Yang, X., Li, J., Lv, N., Liu, F., Wu, J., et al. (2016). The Bacterial Mobile Resistome Transfer Network Connecting the Animal and Human Microbiomes. *Appl Env. Microbiol* 82, 6672–6681. doi:10.1128/AEM.01802-16.

Hultman, J., Rahkila, R., Ali, J., Rousu, J., and Björkroth, K. J. (2015). Meat Processing Plant Microbiome and Contamination Patterns of Cold-Tolerant Bacteria Causing Food Safety and Spoilage Risks in the Manufacture of Vacuum-Packaged Cooked Sausages. *Appl. Environ. Microbiol.* 81, 7088–7097. doi:10.1128/AEM.02228-15.

Jami, E., Israel, A., Kotser, A., and Mizrahi, I. (2013). Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J.* 7, 1069–1079. doi:10.1038/ismej.2013.2.

Jami, E., and Mizrahi, I. (2012). Composition and Similarity of Bovine Rumen Microbiota across Individual Animals. *PLOS ONE* 7, e33306. doi:10.1371/journal.pone.0033306.

Jia, M., Geornaras, I., Belk, K. E., and Yang, H. (2017). 116: Sequence specific removal of Shiga Toxin-Producing Escherichia Coli using the CRISPR-CAS9 System. *Proc. Reciprocal Meat Conf.*, 116.

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., et al. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00459.

Kanwar, N., Scott, H. M., Norby, B., Loneragan, G. H., Vinasco, J., Cottell, J. L., et al. (2014). Impact of treatment strategies on cephalosporin and tetracycline resistance gene quantities in the bovine fecal metagenome. *Sci. Rep.* 4. doi:10.1038/srep05100.

Katoh, K., and Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30, 772–780. doi:10.1093/molbev/mst010.

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., et al. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* 8, 761–763. doi:10.1038/nmeth.1650.

Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinforma. Oxf. Engl.* 28, 3211–3217. doi:10.1093/bioinformatics/bts611.

Kutzer, P., Schulze, C., Engelhardt, A., Wieler, L. H., and Nordhoff, M. (2008). Helcococcus ovis, an Emerging Pathogen in Bovine Valvular Endocarditis. *J. Clin. Microbiol.* 46, 3291–3295. doi:10.1128/JCM.00867-08.

Lakhani, K. R., Boudreau, K. J., Loh, P.-R., Backstrom, L., Baldwin, C., Lonstein, E., et al. (2013). Prize-based contests can provide solutions to computational biology problems. *Nat. Biotechnol.* 31, 108–111. doi:10.1038/nbt.2495.

Lakin, S. M., Dean, C., Noyes, N. R., Dettenwanger, A., Ross, A. S., Doster, E., et al. (2017). MEGARes: an antimicrobial resistance database for high throughput sequencing. *Nucleic Acids Res.* 45, D574–D580. doi:10.1093/nar/gkw1009.

Lange, M., Westermann, P., and Ahring, B. K. (2005). Archaea in protozoa and metazoa. *Appl. Microbiol. Biotechnol.* 66, 465–474. doi:10.1007/s00253-004-1790-4.

Lasken, R. S., and McLean, J. S. (2014). Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat. Rev. Genet.* 15, 577–584. doi:10.1038/nrg3785.

Legendre, P., and Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia* 129, 271–280. doi:10.1007/s004420100716.

Leinonen, R., Sugawara, H., and Shumway, M. (2011). The Sequence Read Archive. *Nucleic Acids Res.* 39, D19–D21. doi:10.1093/nar/gkq1019.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi:10.1093/bioinformatics/btp324.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi:10.1093/bioinformatics/btp352.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., et al. (2010). De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20, 265–272. doi:10.1101/gr.097261.109.

Limam, R. D., Chouari, R., Mazéas, L., Wu, T.-D., Li, T., Grossin-Debattista, J., et al. (2014). Members of the uncultured bacterial candidate division WWE1 are implicated in anaerobic digestion of cellulose. *MicrobiologyOpen* 3, 157–167. doi:10.1002/mbo3.144.

Loman, N. J., Constantinidou, C., Christner, M., Rohde, H., Chan, J. Z.-M., Quick, J., et al. (2013). A Culture-Independent Sequence-Based Metagenomics Approach to the Investigation of an Outbreak of Shiga-Toxigenic Escherichia coli O104:H4. *JAMA* 309, 1502–1510. doi:10.1001/jama.2013.3231.

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* 26, 27663.

Marshall, B. M., Ochieng, D., and Levy, S. B. (2009). Commensals: underappreciated reservoir of antibiotic resistance. *Microbe* 4.5, 231–238.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12. doi:10.14806/ej.17.1.200.

Martínez, J. L., Coque, T. M., and Baquero, F. (2014). What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* 13, 116–123.

Martín-Fernández, J.-A., Hron, K., Templ, M., Filzmoser, P., and Palarea-Albaladejo, J. (2015). Bayesian-multiplicative treatment of count zeros in compositional data sets. *Stat. Model.* 15, 134–158. doi:10.1177/1471082X14535524.

Mayo, B., Rachid, C. T. C. C., Alegría, Á., Leite, A. M. O., Peixoto, R. S., and Delgado, S. (2014). Impact of Next Generation Sequencing Techniques in Food Microbiology. *Curr. Genomics* 15, 293–309. doi:10.2174/1389202915666140616233211.

McKeith, R. O., Gray, G. D., Hale, D. S., Kerth, C. R., Griffin, D. B., Savell, J. W., et al. (2012). National Beef Quality Audit-2011: Harvest-floor assessments of targeted characteristics that affect quality and value of cattle, carcasses, and byproducts,. *J. Anim. Sci.* 90, 5135–5142. doi:10.2527/jas.2012-5477.

McLean, D., Biggs, P., Leblanc-Maridor, M., Hall, R., French, N., Pearce, N., et al. (2014). 0164 Metagenomic detection of bacteria in aerosol samples in animal slaughterhouses to develop exposure profiles for an epidemiological analysis. *Occup. Environ. Med.* 71, 0164.

McMurdie, P. J., and Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLOS Comput Biol* 10, e1003531. doi:10.1371/journal.pcbi.1003531.

Mercier, C. F., Boyer, A., and Coissac, E. (2013). SUMATRA and SUMACLUST: fast and exact comparison and clustering of sequences. Available at: http://metabarcoding.org/sumatra.

Meyer, N. F., Erickson, G. E., Klopfenstein, T. J., Greenquist, M. A., Luebbe, M. K., Williams, P., et al. (2009). Effect of essential oils, tylosin, and monensin on finishing steer performance, carcass characteristics, liver abscesses, ruminal fermentation, and digestibility1. *J. Anim. Sci.* 87, 2346–2354. doi:10.2527/jas.2008-1493.

Miller, J. R., Koren, S., and Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics* 95, 315–327. doi:10.1016/j.ygeno.2010.03.001.

Mir, R. A., Weppelmann, T. A., Elzo, M., Ahn, S., Driver, J. D., and Jeong, K. C. (2016). Colonization of Beef Cattle by Shiga Toxin-Producing Escherichia coli during the First Year of Life: A Cohort Study. *PLOS ONE* 11, e0148518. doi:10.1371/journal.pone.0148518.

Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., Salichos, L., et al. (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome Biol.* 17, 53. doi:10.1186/s13059-016-0917-0.

Nagaraja, T. G., and Chengappa, M. M. (1998). Liver abscesses in feedlot cattle: a review. *J. Anim. Sci.* 76, 287–298. doi:10.2527/1998.761287x.

Nagaraja, T. G., and Lechtenberg, K. F. (2007). Liver Abscesses in Feedlot Cattle. *Vet. Clin. North Am. Food Anim. Pract.* 23, 351–369. doi:10.1016/j.cvfa.2007.05.002.

Nakajima, Y. (1999). Mechanisms of bacterial resistance to macrolide antibiotics. *J. Infect. Chemother. Off. J. Jpn. Soc. Chemother.* 5, 61–74. doi:10.1007/s101569900000.

Nakamura, S., Maeda, N., Miron, I. M., Yoh, M., Izutsu, K., Kataoka, C., et al. (2008). Metagenomic Diagnosis of Bacterial Infections. *Emerg. Infect. Dis.* 14, 1784–1786. doi:10.3201/eid1411.080589.

Nakatsu, G., Li, X., Zhou, H., Sheng, J., Wong, S. H., Wu, W. K. K., et al. (2015). Gut mucosal microbiome across stages of colorectal carcinogenesis. *Nat. Commun.* 6, 8727. doi:10.1038/ncomms9727.

National Antimicrobial Resistance Monitoring System (2010). 2010 NARMS Executive Report.

National Antimicrobial Resistance Monitoring System (2017). NARMS 2015 Integrated Report. Laural, MD: US Department of Health and Human Services Available at: https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM581468.pdf.

Nieminen, T. T., Koskinen, K., Laine, P., Hultman, J., Säde, E., Paulin, L., et al. (2012). Comparison of microbial communities in marinated and unmarinated broiler meat by metagenomics. *Int. J. Food Microbiol.* 157, 142–149. doi:10.1016/j.ijfoodmicro.2012.04.016.

Noyes, N. R., Weinroth, M. E., Parker, J. K., Dean, C. J., Lakin, S. M., Raymond, R. A., et al. (2017). Enrichment allows identification of diverse, rare elements in metagenomic resistome-virulome sequencing. *Microbiome* 5, 142. doi:10.1186/s40168-017-0361-8.

Noyes, N. R., Yang, X., Linke, L. M., Magnuson, R. J., Dettenwanger, A., Cook, S., et al. (2016). Resistome diversity in cattle and the environment decreases during beef production. *eLife* 5, e13195. doi:10.7554/eLife.13195.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi:10.1101/gr.213959.116.

Oakley, B. B., Lillehoj, H. S., Kogut, M. H., Kim, W. K., Maurer, J. J., Pedroso, A., et al. (2014). The chicken gastrointestinal microbiome. *FEMS Microbiol. Lett.* 360, 100–112. doi:10.1111/1574-6968.12608.

Oyaert, M., Cools, P., Breyne, J., Heyvaert, G., Vandewiele, A., Vaneechoutte, M., et al. (2014). Sepsis with an Atopobium-Like Species in a Patient with Fournier's Gangrene. *J. Clin. Microbiol.* 52, 364–366. doi:10.1128/JCM.02310-13.

Parks, D. H., and Beiko, R. G. (2013). Measures of phylogenetic differentiation provide robust and complementary insights into microbial communities. *ISME J.* 7, 173–183. doi:10.1038/ismej.2012.88.

Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013a). Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* 10, 1200–1202. doi:10.1038/nmeth.2658.

Paulson, J. N., Stine, O. C., Bravo, H. C., and Pop, M. (2013c). Robust methods for differential abundance analysis in marker gene surveys. *Nat. Methods* 10, 1200–1202. doi:10.1038/nmeth.2658.

Penadés, J. R., Chen, J., Quiles-Puchalt, N., Carpena, N., and Novick, R. P. (2015). Bacteriophage-mediated spread of bacterial virulence genes. *Curr. Opin. Microbiol.* 23, 171–178. doi:10.1016/j.mib.2014.11.019.

Peng, Y., Leung, H. C. M., Yiu, S. M., and Chin, F. Y. L. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28, 1420–1428. doi:10.1093/bioinformatics/bts174.

Pothakos, V., Stellato, G., Ercolini, D., and Devlieghere, F. (2015). Processing Environment and Ingredients Are Both Sources of Leuconostoc gelidum, Which Emerges as a Major Spoiler in Ready-To-Eat Meals. *Appl. Environ. Microbiol.* 81, 3529–3541. doi:10.1128/AEM.03941-14.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLOS ONE* 5, e9490. doi:10.1371/journal.pone.0009490.

Pukrop, J. R., Campbell, B. T., and Schoonmaker, J. P. (2017). 090 Effect of essential oils or tylosin on performance, liver abscesses, carcass characteristics, and meat quality in feedlot steers. *J. Anim. Sci.* 95, 42–42. doi:10.2527/asasmw.2017.090.

Ransom-Jones, E., Jones, D. L., McCarthy, A. J., and McDonald, J. E. (2012). The Fibrobacteres: an important phylum of cellulose-degrading bacteria. *Microb. Ecol.* 63, 267–281. doi:10.1007/s00248-011-9998-1.

Reti, K. L., Thomas, M. C., Yanke, L. J., Selinger, L. B., and Inglis, G. D. (2013). Effect of antimicrobial growth promoter administration on the intestinal microbiota of beef cattle. *Gut Pathog.* 5, 8. doi:10.1186/1757-4749-5-8.

Rice, W. C., Galyean, M. L., Cox, S. B., Dowd, S. E., and Cole, N. A. (2012). Influence of wet distillers grains diets on beef cattle fecal bacterial community structure. *BMC Microbiol.* 12, 25. doi:10.1186/1471-2180-12-25.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47–e47. doi:10.1093/nar/gkv007.

Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., et al. (2009). Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *Isme J.* 3, 700.

Säde, E., Penttinen, K., Björkroth, J., and Hultman, J. (2017). Exploring lot-to-lot variation in spoilage bacterial communities on commercial modified atmosphere packaged beef. *Food Microbiol.* 62, 147–152. doi:10.1016/j.fm.2016.10.004.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5463–5467.

Scanlan, C. M., and Hathcock, T. L. (1983). Bovine rumenitis - liver abscess complex: a bacteriological review. *Cornell Vet.* 73, 288–297.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi:10.1128/AEM.01541-09.

Schmidt, J. W., Griffin, D., Kuehn, L. A., and Brichta-Harhay, D. M. (2013). Influence of Therapeutic Ceftiofur Treatments of Feedlot Cattle on Fecal and Hide Prevalences of Commensal Escherichia coli Resistant to Expanded-Spectrum Cephalosporins, and Molecular Characterization of Resistant Isolates. *Appl Env. Microbiol* 79, 2273–2283. doi:10.1128/AEM.03592-12.

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18.

Segata, N., Waldron, L., Ballarini, A., Narasimhan, V., Jousson, O., and Huttenhower, C. (2012). Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* 9, 811–814. doi:10.1038/nmeth.2066.

Shahinas, D., Silverman, M., Sittler, T., Chiu, C., Kim, P., Allen-Vercoe, E., et al. (2012). Toward an Understanding of Changes in Diversity Associated with Fecal Microbiome Transplantation Based on 16S rRNA Gene Deep Sequencing. *mBio* 3, e00338-12. doi:10.1128/mBio.00338-12.

Simon, C., and Daniel, R. (2011). Metagenomic Analyses: Past and Future Trends. *Appl. Environ. Microbiol.* 77, 1153–1161. doi:10.1128/AEM.02345-10.

Sneeringer, S., MacDonald, J. M., Key, N., McBride, W. D., and Mathews, K. (2015). Economics of Antibiotic Use in U.S. Livestock Production. Rochester, NY: Economic Research Service, USDA Available at: https://www.ers.usda.gov/webdocs/publications/45485/err-200.pdf?v=0 [Accessed July 9, 2018].

Song, Y. G., Shim, S. G., Kim, K. M., Lee, D.-H., Kim, D.-S., Choi, S.-H., et al. (2014). Profiling of the bacteria responsible for pyogenic liver abscess by 16S rRNA gene pyrosequencing. *J. Microbiol. Seoul Korea* 52, 504–509. doi:10.1007/s12275-014-4241-7.

Soucy, S. M., Huang, J., and Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* 16, 472–482. doi:10.1038/nrg3962.

Stokes, H. W., and Gillings, M. R. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiol. Rev.* 35, 790–819. doi:10.1111/j.1574-6976.2011.00273.x.

Suau, A., Bonnet, R., Sutren, M., Godon, J. J., Gibson, G. R., Collins, M. D., et al. (1999). Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl. Environ. Microbiol.* 65, 4799–4807.

Sulonen, A.-M., Ellonen, P., Almusa, H., Lepistö, M., Eldfors, S., Hannula, S., et al. (2011). Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol.* 12, R94. doi:10.1186/gb-2011-12-9-r94.

Sun, L., Pope, P. B., Eijsink, V. G. H., and Schnürer, A. (2015). Characterization of microbial community structure during continuous anaerobic digestion of straw and cow manure. *Microb. Biotechnol.* 8, 815–827. doi:10.1111/1751-7915.12298.

Tadayon, R. A., Cheema, A. H., and Muhammed, S. I. (1980). Microorganisms associated with abscesses of sheep and goats in the south of Iran. *Am. J. Vet. Res.* 41, 798–802.

Tang, Z.-Z., Chen, G., and Alekseyenko, A. V. (2016). PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. *Bioinformatics* 32, 2618–2625. doi:10.1093/bioinformatics/btw311.

Tessler, M., Neumann, J. S., Afshinnekoo, E., Pineda, M., Hersch, R., Velho, L. F. M., et al. (2017). Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Sci. Rep.* 7, 6589. doi:10.1038/s41598-017-06665-3.

The National Antimicrobial Resistance Monitoring System (2016). The National Antimicrobial Resistance Monitoring System Manual of Laboratory Methods. Available at: https://www.fda.gov/downloads/AnimalVeterinary/SafetyHealth/AntimicrobialResistance/NationalAntimicrobialResistanceMonitoringSystem/UCM528831.pdf [Accessed September 3, 2017].

Thomas, K., Weinroth, M., Parker, J., Metcalf, J., Arthur, T., Schmidt, J., et al. (2018). A Comparision of the Resistome Between Natural an Conventional Retail Ground Beef Productions. in *RMC Annual Meeting Proceedings* (Kanasas City, MI), 132. Available at: http://meatscience.org/docs/default-source/events-and-education/rmc/amsa-2018-71st-rmc-abstracts-7-18-2018-final.pdf?sfvrsn=8ae28fb3_0.

Thomas, M., Webb, M., Ghimire, S., Blair, A., Olson, K., Fenske, G. J., et al. (2017). Metagenomic characterization of the effect of feed additives on the gut microbiome and antibiotic resistome of feedlot cattle. *Sci. Rep.* 7, 12257. doi:10.1038/s41598-017-12481-6.

Thomas, T., Gilbert, J., and Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microb. Inform. Exp.* 2. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3351745/.

Tyson, G. H., Nyirabahizi, E., Crarey, E., Kabera, C., Lam, C., Rice-Trujillo, C., et al. (2018). Prevalence and Antimicrobial Resistance of Enterococci Isolated from Retail Meats in the United States, 2002 to 2014. *Appl. Environ. Microbiol.* 84, e01902-17. doi:10.1128/AEM.01902-17.

Tzeneva, V. A., Heilig, H. G. H. J., van Vliet, W. A., Akkermans, A. D. L., de Vos, W. M., and Smidt, H. (2008). 16S rRNA targeted DGGE fingerprinting of microbial communities. *Methods Mol. Biol. Clifton NJ* 410, 335–349.

USDA–APHIS–VS–CEAH–NAHMS (2013). Feedlot 2011 Part IV: Health and Health Management on U.S. Feedlots with a Capacity of 1,000 or More Head. Fort Collins, CO Available at: https://www.aphis.usda.gov/animal_health/nahms/feedlot/downloads/feedlot2011/Feed11_dr_PartIV.pdf.

USDA-ERS (2013). Organic Production. Available at: https://www.ers.usda.gov/data-products/organic-production.aspx.

Valderrama, W. B., Dudley, E. G., Doores, S., and Cutter, C. N. (2016). Commercially Available Rapid Methods for Detection of Selected Food-borne Pathogens. *Crit. Rev. Food Sci. Nutr.* 56, 1519–1531. doi:10.1080/10408398.2013.775567.

van Dijk, E. L., Jaszczyszyn, Y., and Thermes, C. (2014). Library preparation methods for next-generation sequencing: Tone down the bias. *Exp. Cell Res.* 322, 12–20. doi:10.1016/j.yexcr.2014.01.008.

Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A., and Knight, R. (2013). EMPeror: a tool for visualizing high-throughput microbial community data. *GigaScience* 2. doi:10.1186/2047-217X-2-16.

Vikram, A., Rovira, P., Agga, G. E., Arthur, T. M., Bosilevac, J. M., Wheeler, T. L., et al. (2017). Impact of "Raised without Antibiotics" Beef Cattle Production Practices on Occurrences of Antimicrobial Resistance. *Appl. Environ. Microbiol.* 83, e01682-17. doi:10.1128/AEM.01682-17.

von Wintersdorff, C. J. H., Penders, J., van Niekerk, J. M., Mills, N. D., Majumder, S., van Alphen, L. B., et al. (2016). Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front. Microbiol.* 7. doi:10.3389/fmicb.2016.00173.

Wang, W., Li, C., Li, F., Wang, X., Zhang, X., Liu, T., et al. (2016). Effects of early feeding on the host rumen transcriptome and bacterial diversity in lambs. *Sci. Rep.* 6, 32479. doi:10.1038/srep32479.

Weese, J. S., and Jelinski, M. (2017). Assessment of the Fecal Microbiota in Beef Calves. *J. Vet. Intern. Med.* 31, 176–185. doi:10.1111/jvim.14611.

Weinroth, M. D., Carlson, C. R., Martin, J. N., Metcalf, J. L., Morley, P. S., and Belk, K. E. (2017b). Rapid Communication: 16S ribosomal ribonucleic acid characterization of liver abscesses in feedlot cattle from three states in the United States1. *J. Anim. Sci.* 95, 4520–4525. doi:10.2527/jas2017.1743.

Weinroth, M. D., Yang, X., Rovira, P., Parker, J. K., Morley, P. S., and Belk, K. E. (2018a). Comparing Fecal and Meat Resistomes in U.S. Beef, Pork, and Broiler Production. in (Melbourne, Australia). Available at: http://icomst-proceedings.helsinki.fi/papers/2018_03_07.pdf.

Weinroth, M., Noyes, N., Yang, X., Rovira, P., Doster, E., Dean, C., et al. (2018b). T4-01: Metagenomic Investigation of Antimicrobial-resistance in Beef, Pork, and Broiler Meat. in *International Association of Food Protection Proceedings* (Salt Lake City, UT). Available at: https://iafp.confex.com/iafp/2018/meetingapp.cgi/Paper/18311.

Weinroth, M., Yang, X., Noyes, N. R., Rovira, P., Martin, J. N., Morley, P. S., et al. (2017c). "Foodborne Pathogens and Virulence in the Microbiome of Cattle Grown Naturally verses Conventionally," in *Proc. 63rd International Congress of Meat Science and Technology, Cork, Ireland* (Wageningen, the Netherlands: Wageningen Academic Publishers), 505–507.

Weinroth, Scott, H. M., Norby, B., Loneragan, G. H., Noyes, N. R., Rovira, P., et al. (2018c). Effects of Ceftiofur and Chlortetracycline on the Resistome of Feedlot Cattle. *Appl. Environ. Microbiol.*, AEM.00610-18. doi:10.1128/AEM.00610-18.

Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5, 27. doi:10.1186/s40168-017-0237-y.

White, D. G., Zhao, S., Sudler, R., Ayers, S., Friedman, S., Chen, S., et al. (2001). The Isolation of Antibiotic-Resistant Salmonella from Retail Ground Meats. *N. Engl. J. Med.* 345, 1147–1154. doi:10.1056/NEJMoa010315.

Wichmann, F., Udikovic-Kolic, N., Andrew, S., and Handelsman, J. (2014). Diverse Antibiotic Resistance Genes in Dairy Cow Manure. *mBio* 5, e01017-13. doi:10.1128/mBio.01017-13.

Woerner, D. R. (2012). Beef From Market Cows. Colorado: National Cattlemen's Beef Association Available at: https://www.beefresearch.org/CMDocs/BeefResearch/PE_White_%20Papers/Beef_from_Market_Cows.pdf.

Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15, R46. doi:10.1186/gb-2014-15-3-r46.

World Health Organization (2012). The evolving threat of antimicrobial resistance Options for action. Available at: http://apps.who.int/iris/bitstream/10665/44812/1/9789241503181_eng.pdf [Accessed September 2, 2017].

World Health Organization (2014). Antimicrobial resistance: global report on surveillance. Geneva, Switzerland: WHO Available at: http://apps.who.int/iris/bitstream/handle/10665/112642/?sequence=1.

World Health Organization (2015). Global Action Plan on Antimicrobial Resistance. Geneva, Switzerland Available at: http://apps.who.int/iris/bitstream/10665/193736/1/9789241509763_eng.pdf?ua=1.

World Health Organization (2016). Critically Important Antimicrobials for Human Medicine.

Yang, X., Noyes, N. R., Doster, E., Martin, J. N., Linke, L. M., Magnuson, R. J., et al. (2016). Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within their microbiome in beef production chain. *Appl. Environ. Microbiol.*, AEM.00078-16. doi:10.1128/AEM.00078-16.

Youssef, N. H., Couger, M. B., McCully, A. L., Criado, A. E. G., and Elshahed, M. S. (2015). Assessing the global phylum level diversity within the bacterial domain: A review. *J. Adv. Res.* 6, 269–282. doi:10.1016/j.jare.2014.10.005.

Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30, 614–620. doi:10.1093/bioinformatics/btt593.

Zhang, S., Yin, Y., Jones, M. B., Zhang, Z., Kaiser, B. L. D., Dinsmore, B. A., et al. (2015). Salmonella Serotype Determination Utilizing High-Throughput Genome Sequencing Data. *J. Clin. Microbiol.* 53, 1685–1692. doi:10.1128/JCM.00323-15.

Zolfo, M., Tett, A., Jousson, O., Donati, C., and Segata, N. (2017). MetaMLST: multi-locus strain-level bacterial typing from metagenomic samples. *Nucleic Acids Res.* 45, e7. doi:10.1093/nar/gkw837.

APPENIDX A:

Supplementary Materials for Chapter 2

Sequins and Mock Communities

For normalization on each data set, the following order was performed: raw data was CSS normalized via 'cumNorm' function in MetagenomeSeq (Paulson et al., 2013a), CSS tables were corrected for lane effect, and then based on the mock communities a false positive threshold was set and subtracted from all counts.

Accounting for Lane Effect. In addition to sequencing samples across different lanes and runs of a NovaSEQ 6000, each lane of samples that was sequenced also had two mock communities (ZymoBIOMICS) of known DNA quantity and 2% meta sequins by DNA molecular weight (Either Mix A or B) on each lane. Upon shotgun sequencing, the resulting mock community/Sequin samples were aligned to a FASTA file of the 86 known meta sequin (https://s3.amazonaws.com/sequins/annotations/Metasequins_details.txt) using bwa-mem (Li and Durbin, 2010). Samtools (Li et al., 2009) was then used to covert the SAM files to BAM files, sort and index them and count alignment numbers with 'idxstats.' From the alignment numbers, total Sequins numbers were calculated for each lane and each lane was normalized based on this number to counter lane effect.

Correcting for False Positives in AMR data with a Mock Community. The ZymoBIOMICS mock community, composed of ten known organisms (*Listeria monocytogenes* - 12%, *Pseudomonas aeruginosa* - 12%, *Bacillus subtilis* - 12%, , *Escherichia coli* - 12%, *Salmonella enterica* - 12%, *Lactobacillus fermentum* - 12%, *Enterococcus faecalis* - 12%, , *Staphylococcus aureus* - 12%, *Saccharomyces cerevisiae* - 2%, and *Cryptococcus neoformans*)

was processed in the same way as other samples in the study (library preparation with a custom library bait capture system specific to AMR genes). After sequencing, the mock communities AMR counts were generated in the same way as described with modification to the AMRplusplus pipeline.

To create a list of what AMR genes were undoubtedly present in the 10 organisms included in the mock community, draft complete genomes FASTA files were obtained from NCBI. The bbmap script 'randomreads.sh' was used to fragment the genomes (2x150) with 500,000 fragments per genome. From there, The synthetic reads were aligned to the MEGARes database (Lakin et al., 2017) using bwa-mem (Li and Durbin, 2010). Samtools (Li et al., 2009) was then used to covert the SAM files to BAM files, sort and index them and count alignment numbers with 'idxstats.' The results count numbers were converted to a presence/absence count table and aggregated across all genomes for one composite count for the mock community.

The known counts for the synthetic reads were compared to the CSS normalized counts that were generated through sequencing. The CSS mock community lowest count that also had the presents of a synthetic DNA reads was established as the cutoff point. From there, the CSS normalized cutoff number threshold was subtracted from every count in the count table at the HMM model level. Any value that was below zero was treated as a zero and any HMM model that had zero hits across all samples as a result of this action were removed from the count table.