THE INFLUENCE OF VIRAL AND BACTERIAL COMMUNITY DIVERSITY ON PATHOGENIC *E. COLI* PREVALENCE IN PRE-HARVEST CATTLE

by

Jessica Chopyk

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Plant and Soil Sciences

Summer 2015

THE INFLUENCE OF VIRAL AND BACTERIAL COMMUNITY DIVERSITY ON PATHOGENIC *E. COLI* PREVALENCE IN PRE-HARVEST CATTLE

by

Jessica Chopyk

Approved: _____

K. Eric Wommack, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Blake C. Meyers, Ph.D.
Chair of the Department of Plant and Soil Sciences

Approved: _____

Mark Rieger, Ph.D.
Dean of the College of Agriculture and Natural Resources

Approved: _____

James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education

ACKNOWLEDGMENTS

I would like to foremost thank my advisor, Eric Wommack. It has been a special privilege to work with someone so passionate and sincere about science and scientific integrity. Thank you for giving me the opportunity, first as a lowly intern and than as a masters student. I am forever in your debt. I would also like to thank all the members of my thesis committee for their time, scientific expertise and assistance. Shawn Polson, who has supported, encouraged and helped me throughout my undergraduate and graduate education and Nicole Donofrio, who is an inspiring and passionate role model both in and out of the lecture hall.

To my lab members (current and past), sharing a lab and a workspace with you has something I will never forget. In particular, I want to mention Eric Sakowski, who began this journey with me before I even knew how to pipette and whose patience is an absolute inspiration. Ryan Moore, my partner in STEC CAP crime, whose proficiency in mathematics was instrumental in this thesis. Finally, Daniel Nasko, a man I could not justifiably capture in words, but without him I would not be here today.

Of course, I would like to extend a special thanks to my parents, June and John Chopyk who have stressed the value of education my entire life and have sacrificed tremendously for that education. I am honored and proud to be your (favorite) daughter.

I would also like to acknowledge my funding through the Shiga Toxin-producing *Escherichia coli* Coordinated Agricultural Project (STEC-CAP) grant provided by the United States Department of Agriculture National Institute of Food and Agriculture.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Since 1982, Shiga toxin-producing *Escherichia coli* (STEC) has been an infamous foodborne pathogen causing significant human illness worldwide. The past decades have brought much advancement in food safety practices and industry standards, but STEC still remains prevalent in beef processing with cattle hide implicated as a major source of carcass contamination. To investigate the association between STEC prevalence and the indigenous microbial population of the hide, 16S rRNA bacterial community profiles and viral shotgun metagenomes were created from hide and fecal samples collected from a large commercial feedlot. For the 16S rRNA dataset, fecal OTUs were subtracted from the OTUs found within each hide 16S amplicon library to focus on bacterial populations found exclusively on the hide. Comparative alpha diversity analysis revealed a significant correlation between low bacterial diversity and samples positive for the presence of *E. coli* O157:H7 and/or the non-O157 groups. This trend continued regardless of diversity metric or fecal OTU presence. Beta diversity data revealed differences in the bacterial community composition between the O157 and non-O157 contamination states, with certain OTUs demonstrating significant changes in relative abundance. The exact nature of this relationship remains a mystery, however phage interaction may play a crucial role due to their ability to control the diversity, abundance and genetic composition of their microbial host populations. By dissecting the population ecology of viral groups present in the fecal viral metagenomes we were able to uncover trends in phage host interactions between the feces and the hide and associations between virulent phage and the commensal microbes.

Chapter 1

INTRODUCTION

Shiga toxin-producing *Escherichia coli* (STEC) are zoonotic pathogens that asymptomatically colonize the lower gastrointestinal tracts of cattle and other ruminants (Hancock et al., 1994 and Caprioli et al., 2005). In the United States, STEC cause over two hundred thousand illnesses and approximately 30 deaths each year, generally occurring through the ingestion of contaminated beef products (Scallan et al., 2011). However, STEC is not limited to beef as a means of infection and has been implicated in outbreaks involving sprouts, municipal water, unpasteurized milk and apple cider, and leafy greens. Even person-to-person contact, day care and nursing centers, and recreational water (e.g. swimming pools) have been vehicles of STEC transmission (Tarr et al., 2005).

STEC primarily cause illness in humans through the production of the A_1B_5 family protein, Shiga toxin (Stx1 and Stx2). In host cells Stx binds to the receptor globotriaosylceramide (Gb3) via its pentameric B subunit and is internalized through receptor mediated endocytosis (Sandvig and Van Deurs et al., 1996). It is trafficked by early endosomes to the endoplasmic reticulum where the A subunit unfolds and is inserted into the membrane (Sandvig and Van Deurs et al., 1996). This induces the cell to undergo endoplasmic reticulum-associated protein degradation, which translocates the A subunit to the cytosol where it is folded into an active fragment that can exert toxic effects (i.e. depurination of 60S ribosomal subunit causing protein synthesis inhibition and apoptosis) (Sandvig and Van Deurs et al., 1996, Paton and Paton, 1998 and Keir et al., 2012). If Stx breaches the epithelial barrier and the accompanying vasculature, the toxin can damage Gb3 abundant kidney cells and cause hemolytic uremic syndrome (HUS) (Karmali et al., 1983 and Richardson et al., 1988). HUS is

characterized by haemolytic anaemia, thrombocytopenia and acute renal failure, which can lead to death (Richardson et al., 1988 and Paton and Paton, 1998). Roughly 1/10 to 1/4 of patients will develop HUS after infection, with children being the most susceptible (Karmali et al., 2004). In addition to Shiga toxins, pathogenic STEC possess additional virulence genes, such as the type III secretion system, translocated intimin receptors (*tir*) and intimin (*eae*). These genes enable the bacterium to tightly adhere to colonic epithelial cells, producing characteristic attaching and effacing (A/E) lesions (Paton and Paton, 1998 and Kaper et al., 2004).

STEC are divided into serogroups and serotypes characterized by their somatic lipopolysaccharide (O antigen) and flagella (H antigen), respectively (Nataro et al., 1998). The STEC serotype O157:H7 (*E.coli* O157) is highly virulent and responsible for the majority of hemolytic uremic syndrome (HUS) cases and, as a result, is the most widely characterized and studied serotype of STEC (Riley et al., 1983, Boyce et al., 1995 and Rangle et al 2005). Its infamy in the United States began in 1993 after it claimed the life of four young children after eating contaminated beef patties from a popular fast food chain. Still, 20-50% of infections worldwide are caused by non-O157 serogroups (the "big six"; O26, O111, O103, O121, O45, and O145) with some linked to major outbreaks of HUS in Austria, Germany and Australia (Elliott et al., 2001 Gerber et al., 2002, Brooks et al., 2005 and Hughes et al., 2006). The prevalence of these serogroups in cattle is varying and may depend on factors including, environment (Van Donkersgoed et al., 1999), diet (Callaway et al., 2009), age (Zhao et al., 2013), and seasonality (Barkocy-Gallagher et al., 2003). A global assessment of beef cattle reports over 3 decades determined that *E.coli* O157:H7 presence in cattle ranged between 0.3-19.7% in feedlot cattle and 0.7% -27.3% in grazing cattle, while non-O157 were between 0.7-44.8% in grazing cattle and 4.6-55.9% in feedlot cattle (Hussein, 2007). To further our knowledge of these deadly pathogens this thesis will focus on the bacterial and viral communities of cattle feces and hide and their association with STEC contamination on pre-harvest cattle.

Chapter 2

16S RRNA ASSESSMENT OF BACTERIAL COMMUNITIES ON PRE-HARVEST CATTLE HIDE

2.1 Introduction

Every microbial environment from the cattle gut to the human intestine is an opportunity for STEC to encounter and compete with other microbial populations. These populations host a complex assembly of indigenous microbes, including bacteria, fungi, archaea, and viruses. In particular, commensal bacteria have been shown to mitigate the proliferation of invading pathogens through predation, nutrient competition, and the excretion of antimicrobial compounds (Buffie et al., 2013, Hibbing et al., 2010, Zhao et al., 2013). *E. coli* O157, in particular, thrive within microbial communities demonstrating lower species diversity (Jiang et al., 2002). Specifically, in soil and manure environments, microbial diversity is negatively correlated with the invasion of *E. coli* O157 and *Listeria monocytogenes* (van Elsas et al., 2012, Zhao et al., 2013, Vivant et al., 2013). These studies suggest that indigenous microbial populations interact, often negatively, with pathogen populations. This inherent protection against invasion is especially evident when the diversity of the microbial community is large enough to occupy a broad spectrum of ecological niches, thereby reducing the likelihood that an alien species could gain traction within the novel environment (Hibbing et al., 2010).

To assess community structure, researchers often employ 16S rRNA sequencing, which utilizes the conserved and hypervariable regions of the 16S ribosomal RNA gene to classify sequences into operational taxonomic units (OTUs) or phylotypes. Thousands of environments have been studied via this method and millions of sequences added to online databases, including those from the cattle gut and rumen. From this

work it appears that Firmicutes and Bacteroidetes are the dominant phyla in the cattle gut (Shanks et al., 2011 and Rice et al., 2012) and rumen (Berg Miller et al., 2012 and Jami et al., 2012 and 2014). However, shifts in their relative abundance and other rare community members may occur due to age and fluctuations in diet (Shanks et al., 2011 and McCann et al., 2014). One study also found that butyrate-producing bacteria were negatively associated with STEC shedding and cattle with a higher level of gut bacterial diversity appeared to be associated with lower levels of STEC shedding (Zhao et al., 2013). Therefore, we aim to investigate the microbial community on the cattle hide, a largely unexplored environment.

Cattle hide is a major reservoir of STEC. At a large commercial processing feedlot, 80.7% of sampled cattle hides were positive for at least one of the "big 6" non-O157 and/or O157:H7 (Stromberg et al., 2015). In addition, cattle hide acts as a vehicle of contamination of other animals and the carcass, particularly during transport and slaughter (Bell et al., 1997, Elder et al., 2000, Collis et al., 2004 and Arthur et al., 2007). Hide positive samples have also been shown to be more indicative of carcass contamination than fecal positive samples and hide-to-hide appears to be a larger contamination point than hide-to-pen floor (Keen et al., 2002, Callaway at el., 2009 and Stanford et al., 2011). Interventions focused on the hide, such as dehairing, and washes with water, various chemicals, organic acids and bacteriophage have been developed to stem STEC contamination (Loretz et al., 2011). However, the role of the commensal hide bacterial community in preventing or limiting the prevalence of STEC in cattle is unknown. To determine whether there exists a possible connection between STEC presence and the composition of bacterial communities on cattle hides, we preformed high throughput 16S rRNA sequencing on hide samples exhibiting varying degrees of *E. coli* O157 and non-O157 contamination.

2.2 Methods

2.2.1 Sample Collection

Cattle were sampled over the course of 12 weeks in summer 2013 from a large commercial beef processing facility, both at the feedlot (feces) and at the abattoir (hide). Fresh feces were harvested from two pen floors of a feedlot site 24 hours prior to transport to the harvesting plant where the brisket regions of pre-selected cattle were swabbed with pre-wetted (0.1% sterile peptone water) sponges after stunning and bleeding, but prior to hide removal. Fecal and hide samples were snap-frozen in liquid nitrogen (LN2) within hours of collection and stored at $-80°$. A portion of each hide sample was sent to NeoSeek™(©Neogen, Lansing, MI) for STEC serogroup identification for the serogroups O157, O26, O45, O145, O103, O111 and O121.

2.2.2 Microbial Community Processing

Microbial nucleic acids were extracted from 192 fecal and 182 hide samples taken across the 12-week study via the MO BIO PowerViral DNA/RNA Isolation Kit Tm. Following quantification with the Qubit® Fluorometer, the V3-V4 region of the 16S rRNA gene was amplified using dual-indexed primers (Fadrosh et al., 2014), with an annealing temperature of 52°for 32 cycles. Amplicon reactions (3 µl) were run on an agarose gel to verify amplification. Amplicons were normalized with the Sequal-Prep™Normalization Plates (Invitrogen Inc., CA, USA), pooled, and then sequenced on the Illumina MiSeq platform.

2.2.3 Microbial Community Analysis

Raw Illumina sequences were subjected to quality control and demultiplexing as in Fadrosh et al., 2014. Processed sequences were classified using open-reference OTU picking in Qiime (97% identity) (Caporaso et al., 2010). Briefly, step 1: sequences were clustered against GreenGenes v13.8 (DeSantis et al., 2006). Step 2: sequences that failed to cluster with the reference database were subsequently clustered *de novo*, and the centroids of these clusters were used as reference sequences for the reference

database in step 3. Step 3: sequences that failed the first clustering were clustered against the new reference database generated in step 2. Step 4: The remaining unclassified sequences were clustered *de novo*. All reads were clustered using UCLUST (Edgar, 2010) and aligned with PyNAST (Caporaso et al., 2010). Community richness was evaluated for each sample using the Chao1 estimator (Chao 1984) and Faith's phylogenetic diversity (PD) (Faith 1992). Additionally samples were pooled by various metadata categories (STEC infection status, number of STEC strains, etc.) to examine differences in richness associated with each group. PCoA plots were generated with Emperor (Vazquez-Baeza et al., 2013) using weighted and unweighted UniFrac (Lozupone and Knight 2005) distance to evaluate large-scale structural shifts associated with each metadata category. Taxonomic profiles for each metadata group were analyzed with Kruskal-Wallis one-way analysis of variance and paired Mann-Whitney U tests to identify taxa at significantly higher or lower abundance, and significant OTUs were visualized with Cytoscape (Shannon et al., 2003).

2.3 Results

2.3.1 Sequencing and STEC Identification

After assessing the data for poor quality and removing singletons, 1,688,954 high-quality 16S rRNA gene sequences were recovered from hide bacterial samples. From these sequences, 97,050 OTUs were classified across the dataset. Each sample provided an average of 9,877 sequences. In ranking the abundance of the OTUs the 80%, 90%, 95%, and 99% of the abundance occurred in the top 225, 554, 1165, and 4062 OTUs, respectively (Fig. 2.1). Feces are often found on the hide and likely influenced the density of bacteria in hide swab samples. To focus on bacterial populations found only on the hide and not in feces, OTUs identified within fecal samples were subtracted from the OTUs found within each hide 16S amplicon library. After removal of sequences identified as belonging to fecal bacterial populations, 135,202 total sequences remained. Within this collection of sequences a total of 2,709 hide-only OTUs (97% identity) were identified across the 182 hide swab libraries. For serogroup typing, the NeoSeek™assay

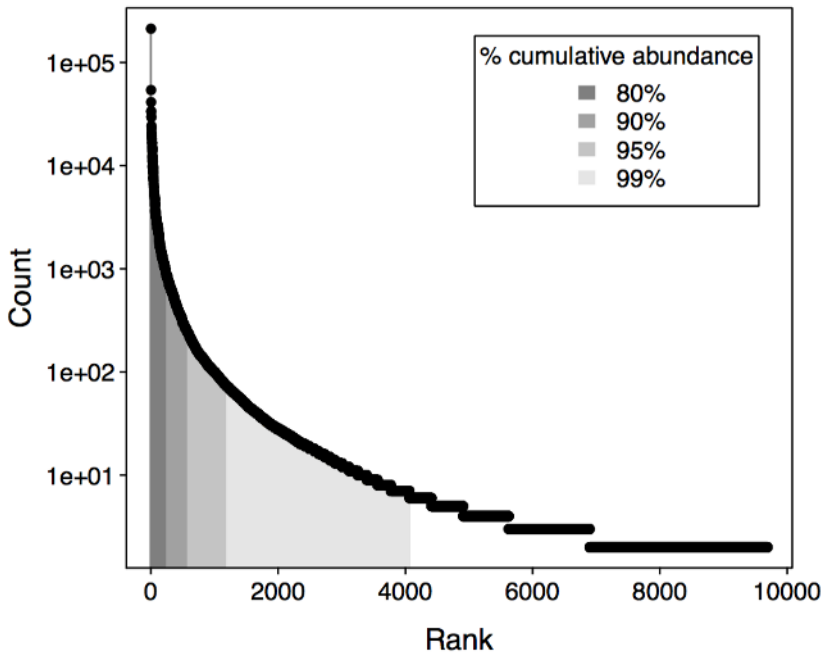


Figure 2.1: Rank abundance of OTUs. Rank for OTUs occurring at least twice. The shade bars under the curve represent the percent abundance contained in that section of the curve. Eighty percent, 90%, 95%, and 99% of the abundance was contained in the top 225, 554,1165, and 4062 OTUs, respectively.

indicated that of the 182 hide samples, 28 were negative for STEC, 32 were positive for *E. coli* O157, 55 were positive for at least one non-O157 serogroup, and 67 were positive for *E. coli* O157 and at least one other non-O157 serogroup (both) (Stromberg et al., 2015).

2.3.2 Bacterial Community Taxonomy

The principle bacterial families identified in the hide swab samples were Corynebacteriaceae, Ruminococcaceae, Lachnospiraceae, Clostridiaceae, and of these, the Corynebacteriaceae was the largest family regardless of contamination status (Fig. 2.2). When the fecal OTUs were subtracted the dominant bacterial families in the hide-only dataset were Moraxellaceae, Staphylococcaceae, Streptococcaceae and Pasteurellaceae, and of

these, the Staphylococcaceae was the dominant family in the STEC positive samples, O157 (39%), non-O157 (33%), and both (34%), while the Moraxellaceae was the largest in the STEC negative sample (33%) (Fig. 2.3). Streptococcaceae populations comprised a lower relative abundance in *E. coli* O157 (12%) than in the non-O157 (20%) and STEC negative samples (21%).

2.3.3 Alpha Diversity Metrics by Contamination Status

Alpha diversity was lower in cattle hide samples grouped by contamination status when STEC was detected, with the lowest diversity appearing in samples that tested positive for both O157 and non-O157 (Fig. 2.4A). The trend continued regardless of whether the diversity was measured by phylogenetic distance, the Chao1 species estimator or observed OTUs (Fig. 2.4A). Among the hide-only dataset this trend was maintained as alpha diversity was lower for STEC positive samples when compared to samples without detectable STEC (Fig. 2.5A). Additionally, samples were pooled based on number of STEC serogroups identified, regardless of whether they were O157 or non-O157 positive. When the fecal OTUs were present alpha diversity decreased significantly between STEC negative samples, and samples containing one, two, or three or more serogroups, for each diversity metric tested (Fig. 2.4B). This trend was maintained in the hide-only dataset (Fig. 2.5B).

2.3.4 Beta Diversity Metrics by Contamination Status

Beta diversity was compared between contamination status (i.e. samples containing either O157, non-O157 or both; Fig. 2.6A) and the number of serogroups identified within samples (i.e. 1, 2, or 3 or more serogroups; Fig. 2.6B) using a three-dimensional principal coordinate analysis with and without fecal OTUs. The comparison of contamination states with fecal OTUs showed clustering not only between STEC negative and positive samples, but also between O157 and non-O157 groups (Fig. 2.6A). Along PC1, which explains 36.49% of the difference, STEC positive samples grouped away from STEC negative samples, while along PC2, 35.57%

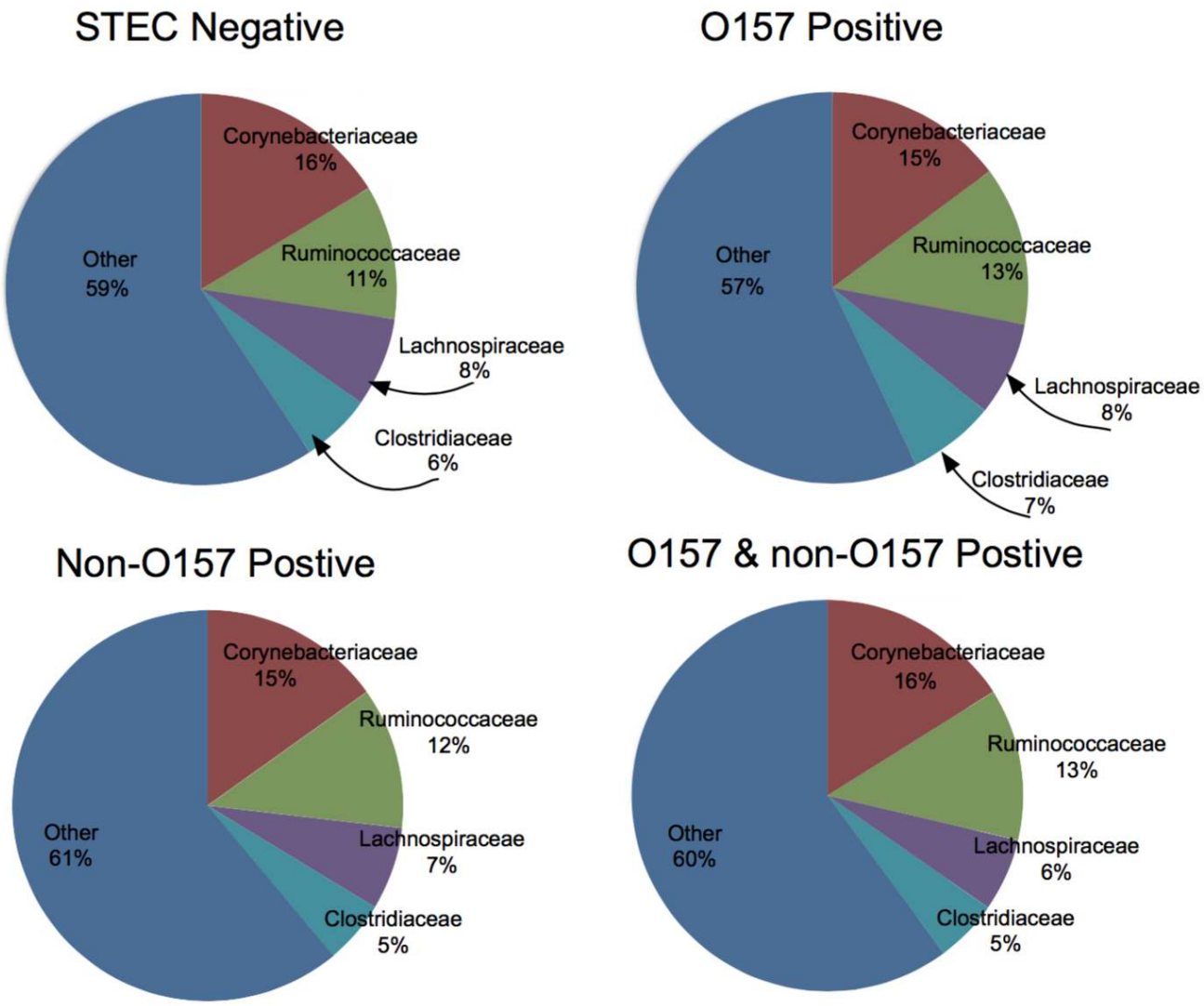


Figure 2.2: Taxonomic distribution of bacterial families from hide samples according to contamination status with all OTUs. Relative abundance charts were based on distributions of bacterial families as a percentage of the total number of classified 16S sequences. Samples were pooled based on STEC contamination status (STEC negative, O157 positive, non-O157 positive, and both).

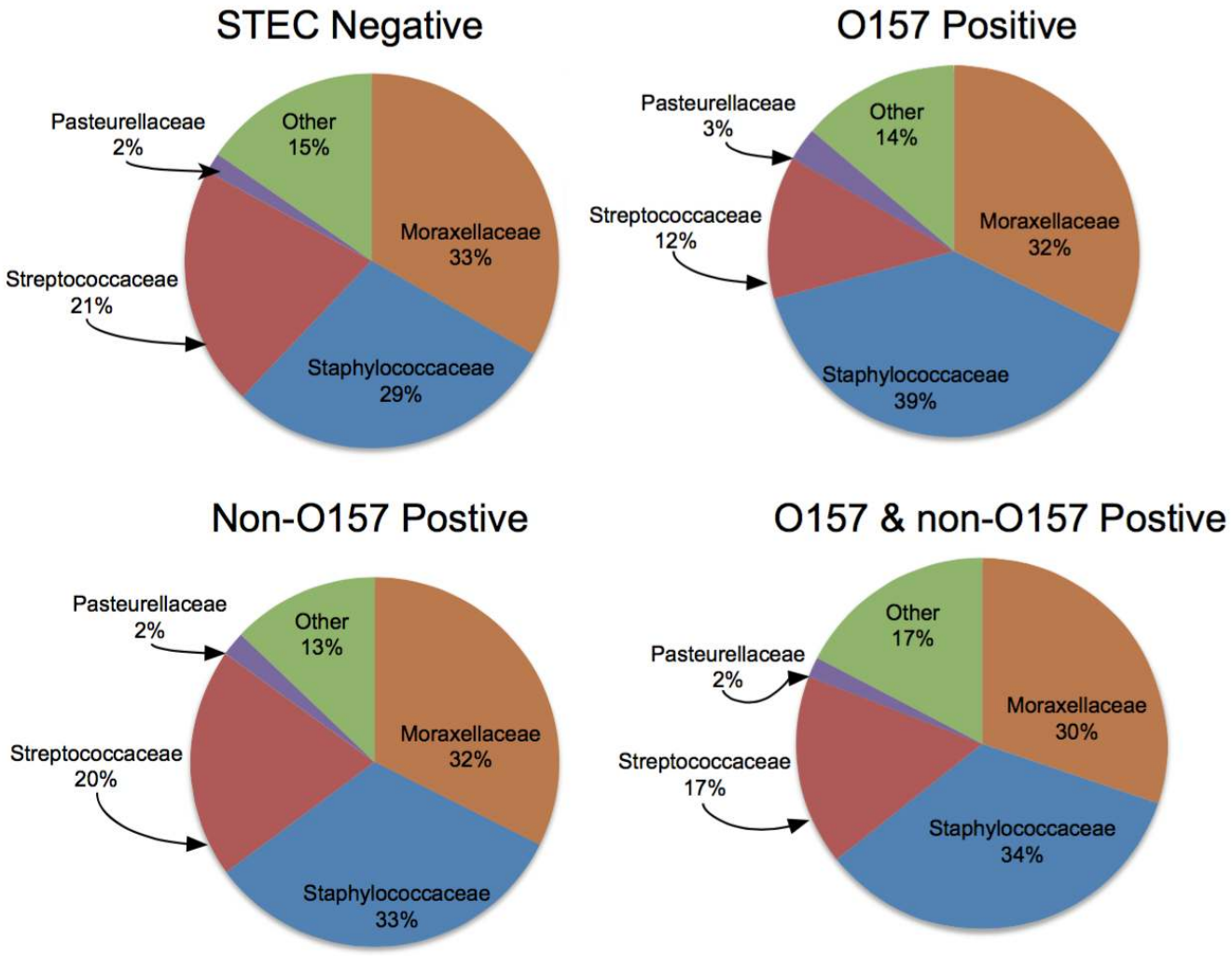


Figure 2.3: Taxonomic distribution of bacterial families from hide samples according to contamination status with hide-only OTUs. OTUs classified from the fecal 16S rRNA survey were subtracted to create the hide-only category. Relative abundance charts were based on distributions of bacterial families as a percentage of the total number of classified 16S sequences. Samples were pooled based on STEC contamination status (STEC negative, O157 positive, non-O157 positive, and both).

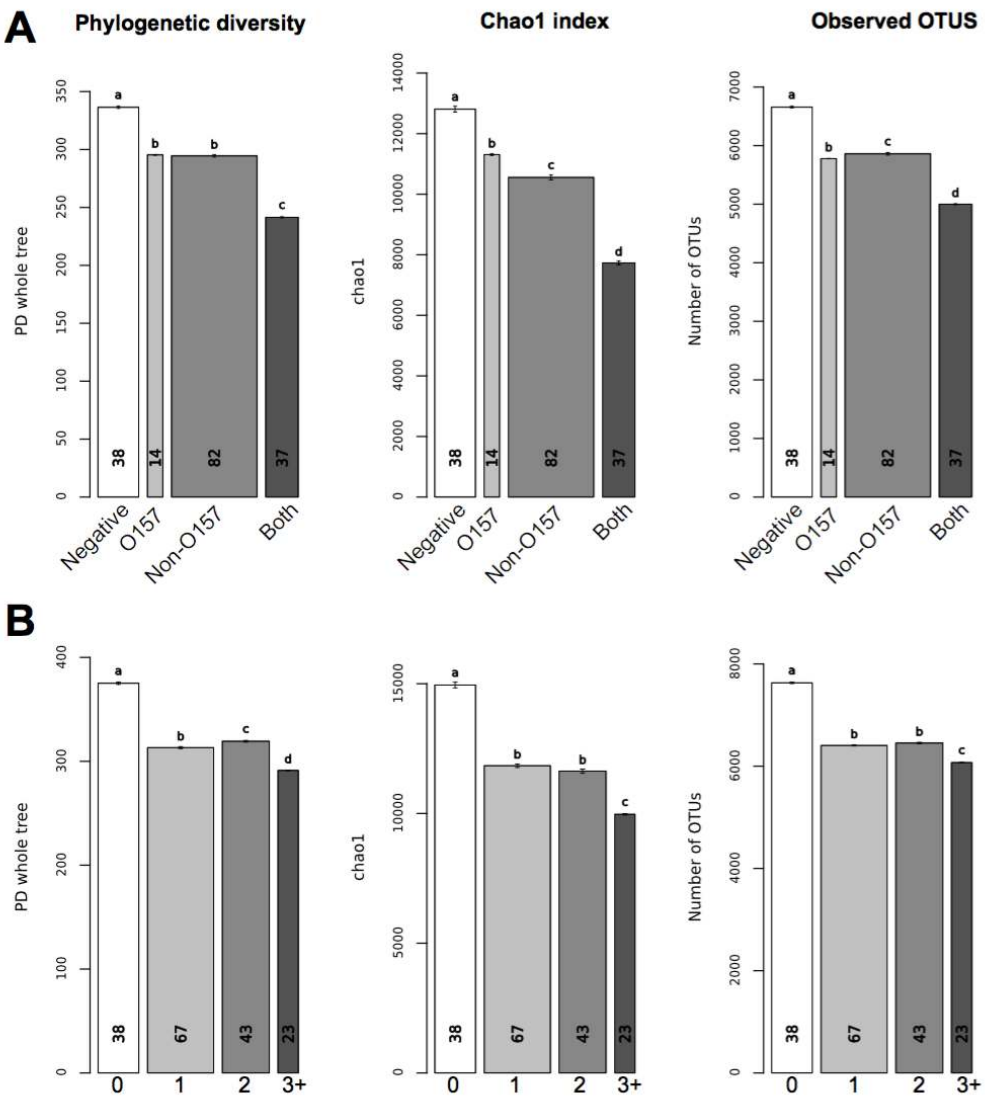


Figure 2.4: Alpha diversity metrics for samples pooled by contamination status with all OTUs. Jackknifed alpha diversity metrics (phylogenetic distance, chao1 estimator, and observed OTUs) were calculated with Qiime's alpha_rarefaction.py script. Phylogenetic distance, the chao1 index, and observed OTUs were calculated for ten jackknifes at a depth of 95% of the reads contained in the smallest metadata group. Error bars represent the standard error of the jackknife estimate. The width of the bar shows the number of sequences contained in each metadata category, while the small numbers inside the bars indicate the number of samples in each category. Within each chart, bars with different letters were significantly different at an alpha level of 0.05 with Bonferonni correction as determined by the Mann-Whitney U test. In (A), samples were pooled by STEC contamination status into four groups: STEC negative, *E. coli* O157 positive, non-O157 positive and positive for both *E. coli* O157 and any of the tested non-O157 serogroups, whereas in (B) samples were pooled by number or STEC serogroups present in the sample

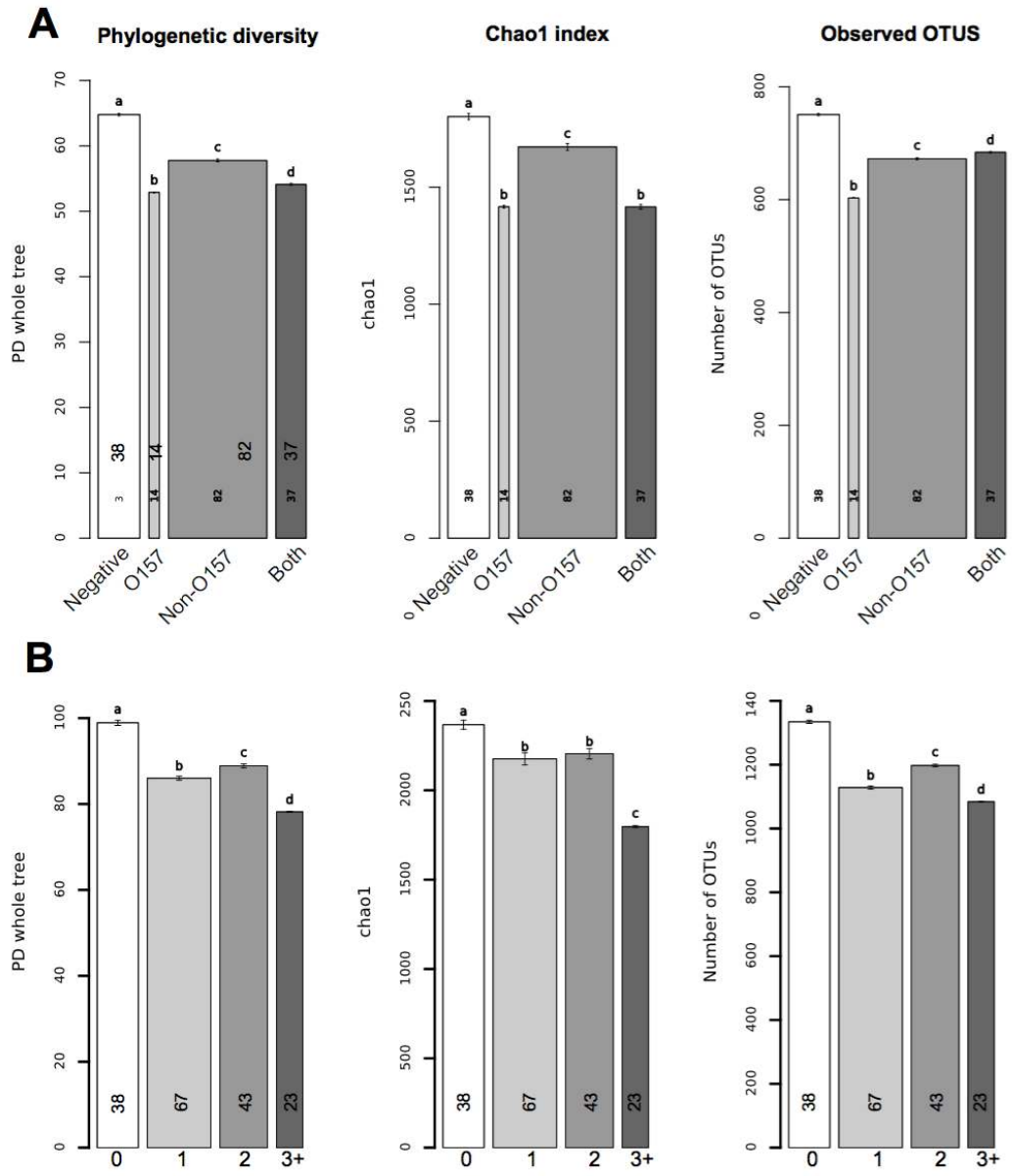


Figure 2.5: Alpha diversity metrics for samples pooled by contamination status with hide-only OTUs. Alpha diversity retested after fecal OTUs were removed. Error bars represent the standard error of the jackknife estimate. The width of the bar shows the number of sequences contained in each metadata category, while the small numbers inside the bars indicate the number of samples in each category. Within each chart, bars with different letters were significantly different at an alpha level of 0.05 with Bonferonni correction as determined by the Mann-Whitney test. In (A), samples were pooled by STEC contamination status into four groups: STEC negative, *E. coli* O157 positive, non-O157 positive and positive for both *E. coli* O157 and any of the tested non-O157 serogroups, whereas in (B) samples were pooled by number or STEC serogroups present in the sample.

of the difference, the samples positive for O157 clustered away from the other categories. When grouped by number of STEC strains (Fig. 2.6B) the negative samples (0 strains) and positive samples (1+ strains) again clustered away from each other along PC1 (38.12% of the difference), while those with one STEC strain had the same partitioning as O157 along PC2 (31.42% of the difference) as samples in the plot grouped by contamination status (Fig. 2.6A). Within the one strain group there were 53 samples (79.1%) that contained a non-O157 serogroup compared to 14 samples (20.9%) with O157. Thus, the O157 group in Fig. 2.6A, is analogous to the one strain group in Fig. 2.6B, with the non-O157 samples removed. Because both groups show similar partitioning, it appears that O157 is driving the separation. When the fecal OTUs were removed (Fig.2.7), the weighted Unifrac measure had to be employed. Along PC1, which explains 55.36% of the difference, STEC categories grouped away from each other (Fig. 2.7A). When grouped by number of STEC serogroups identified the same portioning can be seen as in Fig. 2.6B. Along PC1 (72.72% of the difference) the samples with one serogroup (O157) and the samples negative for all STEC (0 strains) clustered away from samples with 1+ serogroups.

2.3.5 OTUs demonstrating significant change with STEC contamination status

OTUs at significantly higher or lower relative abundance depending on contamination status were identified when fecal OTUs were present and absent (Fig. 2.8). No OTUs were identified as occurring at a higher relative abundance in STEC positive samples in either dataset (Fig. 2.8). Two OTUs genera that demonstrated higher relative abundance in the hide-only STEC negative samples belonged to the genera *Streptococcus* and *Solibacillus*, with the *Streptococcus* OTU #532232 being lower in all STEC positive categories and *Solibacillus* OTU #606419 being significantly lower in non-O157 infected samples (Fig. 2.8B). Three additional OTUs, *Staphylococcus* (OTU #630141) and *Streptococcus* (OTU #270572 and #151098) were also lower in non-O157 infected samples. Two *Streptococcus* OTUs (OTU #4441855 and #439035)

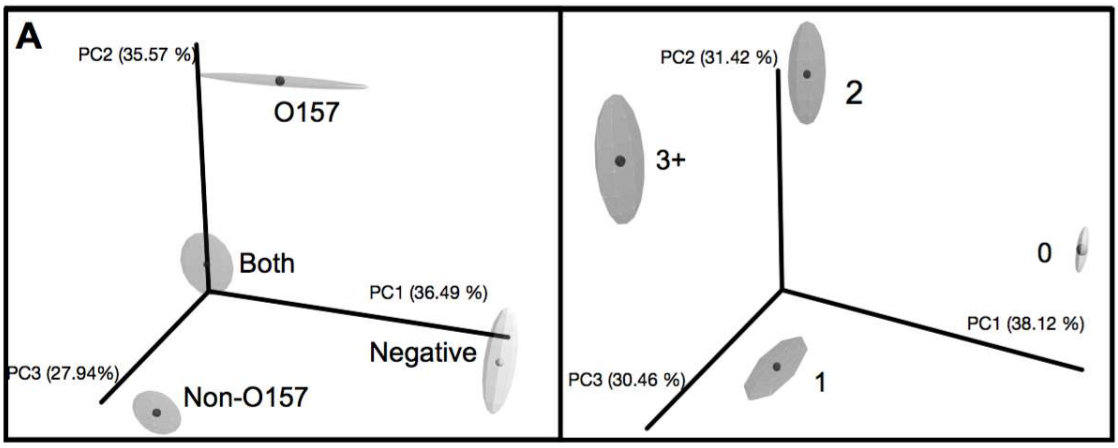


Figure 2.6: Beta diversity comparison of the bacterial communities by contamination status and number of STEC serogroups with all OTUs. Jackknifed principal coordinate analysis using the unweighted UniFrac metric was performed (100 jackknifes) on all hide sequences. Sequences were pooled by contamination state (A), *E. coli* O157 positive, non-O157 positive, positive for O157 and at least one non-O157 STEC positive group and samples negative for all STEC serogroups tested and by number of STEC serogroup detected (B). Percentage of variation explained by principal coordinate shown on each of the axis.

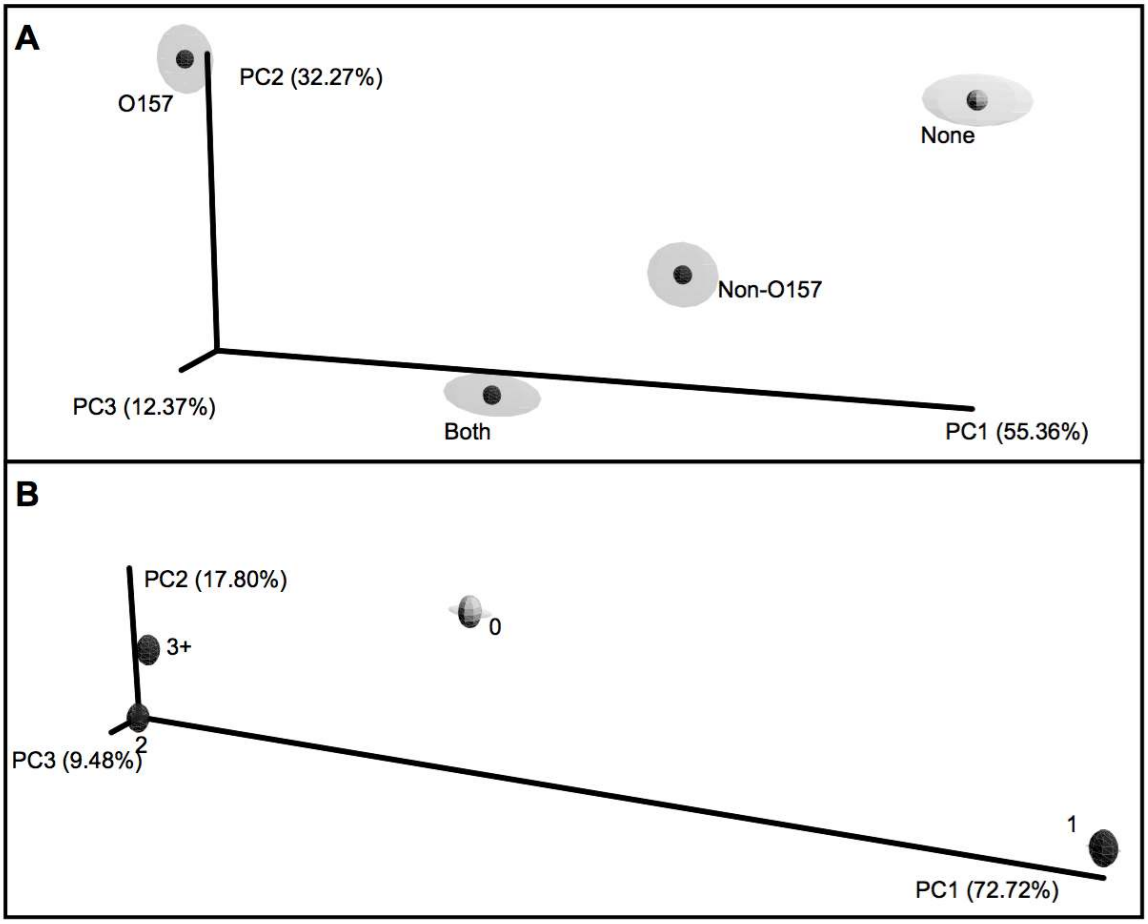


Figure 2.7: Beta diversity comparison of the bacterial communities by contamination status and number of STEC serogroups with hide-only OTUs. Jackknifed principal coordinate analysis using the weighted UniFrac metric was performed (100 jackknifes) on hide-only OTUs by STEC contamination state (A) and by number of STEC serogroup detected (B). Percentage of variation explained by principal coordinate shown on each of the axis.

were also lower in the both category and one from the O157 positive category (OTU #532232). The other O157 positive OTUs that were lower were *Deinococcus* (OTU #359868) and two additional *Streptococci* (OTU #288249 and #322775). These OTUs (OTU #359868, #288249 and #32277) were all also lower in the both category along with *Mannheimia* (OTU #440422), *Enhydrobacter* (OTU #1114747), *Moraxella* (OTU #8281, #334337, #367879), *Alloiococcus* (OTU #15238), and *Acinetobacter* (OTU #4449979) (Fig. 2.8B).

The OTUs *Brevibacterium* (OTU #73538) and *Ruminococcus* (OTU #532187) and Nocardioidaceae (OTU #528213) were at higher relative abundance in STEC negative samples when fecal OTUs were present. All three were also lower in non-O157, while *Brevibacterium* (OTU #73538) and *Ruminococcus* (OTU #532187) were significantly lower in non-O157 and in the both group (Fig. 2.8A). Thermoactinomycetaceae (OTU #4400372), *Dietzia* (OTU #1126467), Clostridales (OTU #110678), and *Ruminococcus* (OTU #333948) were also lower in the non-O157 and the both category. For the O157 positive category the remaining reduced OTUs were Mogibacteriaceae (OTU #4295063), *Selenomonas* (OTU #311471), Rumiococcaceae (OTU #515308), *Streptococcus* (OTU #532232) and Clostridales (OTU #560906), with all but Clostridales (OTU #560906) shared with both. The remaining OTUs that were significantly lower in the both category were largely gram-positive bacteria Lachnospiraceae (OTU #299184 and #294440), *Jeotgalicoccus* (OTU #4404401), *Fusobacterium* (OTU #2987631), *Trichococcus* (OTU #4403115), Clostridiales (OTU #4482650), *Corynebacteria* (OTU #412872, #235898, and #103606), *Yaniella* (OTU #115315), Clostridiaceae (OTU #4383953), *Butyrivibrio* (OTU #13983) and SMB53 (OTU #180516).

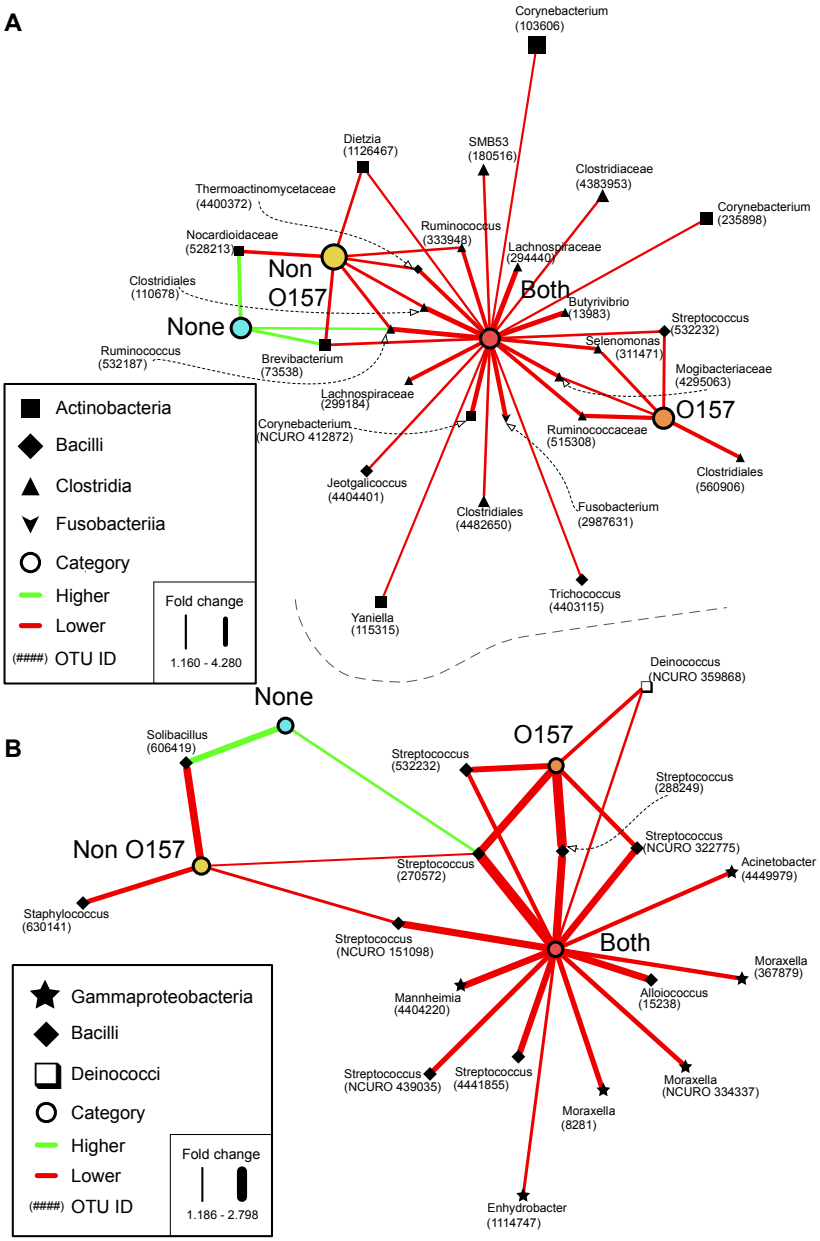


Figure 2.8: OTUs with significantly different relative abundance according to contamination status. A) All OTUs. B) Hide-only OTUs. White circles indicate contamination status: STEC negative; blue, O157 only; orange, non-O157 only; yellow, both non-O157 and O157; red. Black nodes are OTUs with significantly higher or lower relative abundance. The shape of OTU nodes is based on class: Bacilli; Clostridia; Actinobacteria; Gammaproteobacteria; Deinococci; Fusobacteria. Labels indicate the genus of the OTU. A red line represents lower average relative abundance. Green lines indicate higher relative abundance. Edge width is proportional to fold-change, with a thicker line representing a higher shift in abundance. When genus information was not available family or order is indicated.

Chapter 3

METAGENOMICS OF BACTERIOPHAGE IN PRE-HARVEST CATTLE FECES

3.1 Introduction

From our research on the bacterial communities of the cattle hide we know that the alpha diversity decreases in the presence of STEC and the community composition changes depending on contamination state. In order to further understand this relationship it is important to look at the co-occurring bacteriophage, an understudied aspect of microbial ecology in STEC research. In all known environments phage are at the heart of a complex microbial ecosystem. At each beat, phage drive the infection and lysis of host bacterial cells, propagating nutrient and biogeochemical cycles and influencing the genetic architecture of their host (Clokie et al., 2011). However, dissecting the population ecology of viral groups has proven to be technically difficult as there is no gene, akin to 16S rRNA in bacteria, which can be used to examine the diversity and distribution of viral populations in nature.

One way around this limitation has been shotgun viral metagenomics, a cultivation-independent technique whereby purified viral nucleic acids from natural samples are randomly sequenced to create large metagenomic datasets known as viromes. This method has been employed to characterize the genetic diversity of environmental phage from a wide variety of environments, including, oceans, estuaries, soils, hydrothermal vents, hot springs, and organismal substrates (Breitbart et al., 2004a, Angly et al., 2006, Bench et al., 2007, Fierer et al., 2007, Schoenfeld et al., 2008 and Anderson et al., 2014). Despite the fact that over 60% of predicted open reading frames (ORFs) have no homologous representatives among known references, these viomes have revealed clues about the biology and ecology of environmental phage (Breitbart et al.,

2002). For instance, studies on human adult feces have predicted roughly 1200-2000 viral genotypes, whereas feces from infants had only 8 viral genotypes (Breibart et al., 2003 and 2008). Siphoviruses were found to be the dominant family of Caudovirales in the fecal viromes of horses, sea lions and humans (Breibart et al., 2003, Cann et al., 2005 and Li et al., 2011) Although no cattle fecal viral metagenome is currently available, cattle rumen has been shown to house viral communities that are incredibly diverse and unique, with an estimated 28,000 different viral genotypes reported in some individuals (Berg Miller et al., 2012 and Ross et al., 2013).

One of the tools that have emerged from viral metagenomics and PCR-based approaches, is the use of information proteins to explore a subset of phage diversity including ribonucleotide reductases (Dwivedi et al., 2013 and Sakowski et al., 2015), T4 vertex portal protein and major capsid protein (File et al., 2005 and Short et al., 2005), and polymerase A (Breitbart et al., 2004b, Labont et al., 2009, and Schmidt et al., 2014). For this study we sought to employ two of these marker genes, polymerase A and ribonucleotide reductase.

Family A polymerase (PolA) has proven to be a suitable marker gene for phylogenetic characterization and diversity analyses, appearing in 25% of all known dsDNA phage genomes (Wommack et al., 2015), particularly tailed phages (i.e. Podoviridae and Siphoviridae). The primary function of PolA in bacteria, bacterial eukaryotic organelles, and bacteriophage is DNA repair and lagging strand synthesis. Within bacteriophage the polymerase gene lacks the 5'-3' exonuclease domain found in bacteria, making it the chief polymerase for phage genome replication (Doublie et al., 1998). Using PCR based approaches PolA genes in T7-like podoviruses have been identified across multiple environments and, in some cases, exhibiting environmental specificity (Breitbart et al., 2004a, Labont et al., 2009, Schmidt et al., 2014). It has also been reported that an amino acid mutation at position Phe762 relative to *E. coli* introduces biochemical changes within the polymerase that may be indicative of lifestyle (Tabor and Richardson, 1987, Tabor and Richardson, 1995 and Suzuki et al., 2000). A Phe762

to tyrosine substitution (Phe762Tyr) was shown to produce a faster less accurate polymerase, characteristic of virulent phage. Conversely, a Phe762 to leucine mutation (Phe762Leu) produced a slower more accurate polymerase suitable to a temperate lifestyle (Schmidt et al., 2014).

Ribonucleotide reductase (RNR), another prolific maker, is a core housekeeping gene responsible for the formation of deoxyribonucleotides from ribonucleotides, a key step in DNA synthesis (Jordan and Reichard, 1998 and Nordlund et al., 2006). RNRs appear in roughly 17% of dsDNA-tailed phage, particularly virulent Myoviruses and Siphoviruses (Wommack et al., 2015). Analogous to PolA, RNRs can be biologically predictive due to their reactivity with O_2. Class I RNRs are O2-dependent, Class II RNRs are O_2-independent, but need adenosylcobalamin (vitamin B12), and Class III RNRs are O_2-sensitive (Jordan and Reichard, 1998 and Nordlund et al., 2006). Therefore, we can leverage the information from these two vital proteins to assess virus biology and ecology within our viromes.

Using a metgenomic approach on isolated viral DNA from 11 fecal samples we sought to survey the first cattle fecal viromes to determine how they vary among individuals and correlate that with existing 16S rRNA sequencing and STEC metadata. Additionally, we aimed to explore connections between the fecal viromes and what we have uncovered about the diversity and composition of commensal bacterial communities on the hide.

3.2 Methods

3.2.1 Sampling and STEC identification

Over a 12-week period (June to August 2013) fecal samples from pen floors were collected from a large commercial cattle feedlot. Samples were snap frozen and a portion sent for microbial nucleic acid extraction and STEC prevalence detection. For culture based detection of STEC serogroups a collaborating lab (Dewsbury et al., 2015) enriched 2 grams of each fecal sample in *E. coli* broth and an aliquot of 980 µl was added to serogroup-specific IMS beads (Abraxis®, Warminister, PA), for

the STEC serogroups of interest (O26, O45, O103, O111, O121, O145, and O157). The IMS suspensions were spread on MacConkey agar with cefixime and potassium tellurite for O157 and/or modified Poss (Poss et al., 2008) for the non-O157. Following incubation for 24 h, 6 colonies were chosen from each plate and streaked on blood agar plates for another 24 h incubation at 37°. Colonies positive for latex agglutination and indole production on the O157 plates were used in a mulpliex PCR for O157 genes: fliC (encodes the *E. coli* flagellum), rfbE (encodes the *E. coli* O157 serotype), ehxA (enterohemolysin), stx1, stx2 (Bai et al., 2010). Non-O157 colonies were PCR tested for serogroup specific genes (O26, O45, O103, O111, O121, and O145; Paddock et al., 2012) and, if positive, virulence genes (eae, stx1, stx2) (Bai et al., 2012).

3.2.2 Viral Concentrate Construction and Enumeration

An adaption of the $FeCl_3$ method (John et al., 2011) was used to create cell-free viral concentrates. Fecal samples (0.75g) at varying states of STEC contamination were diluted with 50 mL of phosphate buffered saline and gently shaken for 1 h. Diluted samples were filtered with a 0.22 µm polycarbonate filter and spiked with 50 µl of $FeCl_3$. Following incubation at room temperature for 1 h, the FeCl3 flocculate was filtered onto a 1.0 µm polycarbonate filter. Phage within the Fe precipitates were resuspended with 500 µl of oxalic acid buffer. Resuspended samples were digested with 2 h DNAse incubation to eliminate free cellular DNA and then 0.22 µm filtered again. A 16S rRNA PCR was performed to ensure all phage concentrates were free of cellular DNA. Outlined protocol is depicted in Appendix A.

Viral particles were enumerated via epifluorescence microscopy (Winget et al., 2009) with 20l of viral concentrate. Pictures were taken at 1000x magnification with an Olympus BX61 microscope under 100x UPlanFI oil objective and counted with iVision viral test software. Phage morphology was visualized for one sample via transmission electron microscopy.

3.2.3 Viral metagenome sequencing and analysis

The phenol-chloride "crack" method (as described in Appendix B) was used for DNA isolation. DNA was quantified with the Qubit® Fluorometer and roughly 20 ng of cell-free viral DNA from each sample were sequenced with paired-end Illumina HiSeq 2500 (2 x 151 cycles, 1 lane). Fecal 16S rRNA libraries were processed as described in Chapter 1.

SPAdes version 3.5.0 with the BayesHammer read correction module was used to assemble each of the eleven viromes (Bankevich et al., 2012). To account for the variability in coverage inherent in a metagenomic library, the SPAdes was run in single cell mode with kmer sizes of 21, 27, 33, and 41. Assemblies were input into the metagenomic analysis web-tool VIROME (Wommack et al., 2012). Phage taxonomy, top functional proteins and virulence genes were compared.

ORFs were predicted in each virome using MetaGeneMark (Zhu et al., 2010), queried against a database of RNR and PolA UniRef90 clusters with an e-value cutoff of $1e^{-5}$ (Altschul et al., 1990 and Suzek et al., 2007), and sorted based on length (200aa) and NCBI's Conserved Domain BLAST online tool (Marchler-Bauer et al., 2011). PolAs were aligned with MAFFT using the FFT-NS-i 1000 algorithm and region with conserved residues (S582-Y766 relative to *E.coli*) was selected and clustered at 75% using the furthest neighbor algorithm in mothur (Schloss et al., 2009). Cluster representatives were than used to generate an unrooted maximum likelihood tree with 100 bootstrap replicates with PhyML via Geneious 6.0.5 (Guindon et al., 2003 and Katoh et al., 2002). This procedure was repeated for RNR classes I, II and III with selected regions L264-D581, L492-and A741, respectively. Nodes were annotated with tailed phage BLASTp assignment with an e-value cutoff of $1e^{-20}$ (Altschul et al., 1990).

3.3 Results

3.3.1 Viral abundance and morphology

Free virus particles were successfully isolated from 11 fecal samples at varying states of STEC contamination. STEC negative samples had the greatest overall abundance of viral particles at an average of 1.07 x 10^{10} viruses g-1, followed by samples positive for *E.coli* O157 and any of the non-O157 serogroups at 7.36 x 10^{9} viruses g-1. The *E.coli* O157 and non-O157 positive samples had the least amount at 7.00 x 10^{9} and 6.70 x 1^{9} viruses g-1, respectively (Fig. 3.1). Due to the limited sampling depth, there was only a limited statistical power, thus, none had a p value less than 0.05. (STEC negative samples to O157 positive, p value 0.0595, and STEC negative to non-O157, p value 0.08536). Additionally, all three families of the dsDNA-tailed phages, Caudovirales, appeared to be present in transmission electron microscopy images of a non-O157 positive fecal sample (Fig. 3.2).

3.3.2 Virome sequencing and assembly

The MiSeq Illumina sequencing and SPAdes assembly yield is described in Table 3.1. Due to its low sequencing yield, sample 141 was not used for the remaining virome analyses.

3.3.3 Viral and bacterial taxonomy

The dominant bacterial OTUs in the fecal 16S rRNA libraries that corresponded to the viromes were Firmicutes followed by Bacteroidetes, regardless of STEC contamination state (Fig. 3.3A). For STEC negative samples 123 and 130 Firmicutes and Bacteroidetes were 68.71% and 12.29% and 83.83% and 35.09%, respectively. The O157 positives samples 111, 100, 140 and 117 were 62.07% and 28.84%, 83.96% and 9.84%, 73.79% and 12.29%, and 52.50% and 29.10% for Firmicutes and Bacteroidetes, respectively. For the non-O157 positive sample 144 Firmicutes was 62.98% and Bacteroidetes was 25.80% and sample 141 Firmicutes was 64.58% and Bacteroidetes was

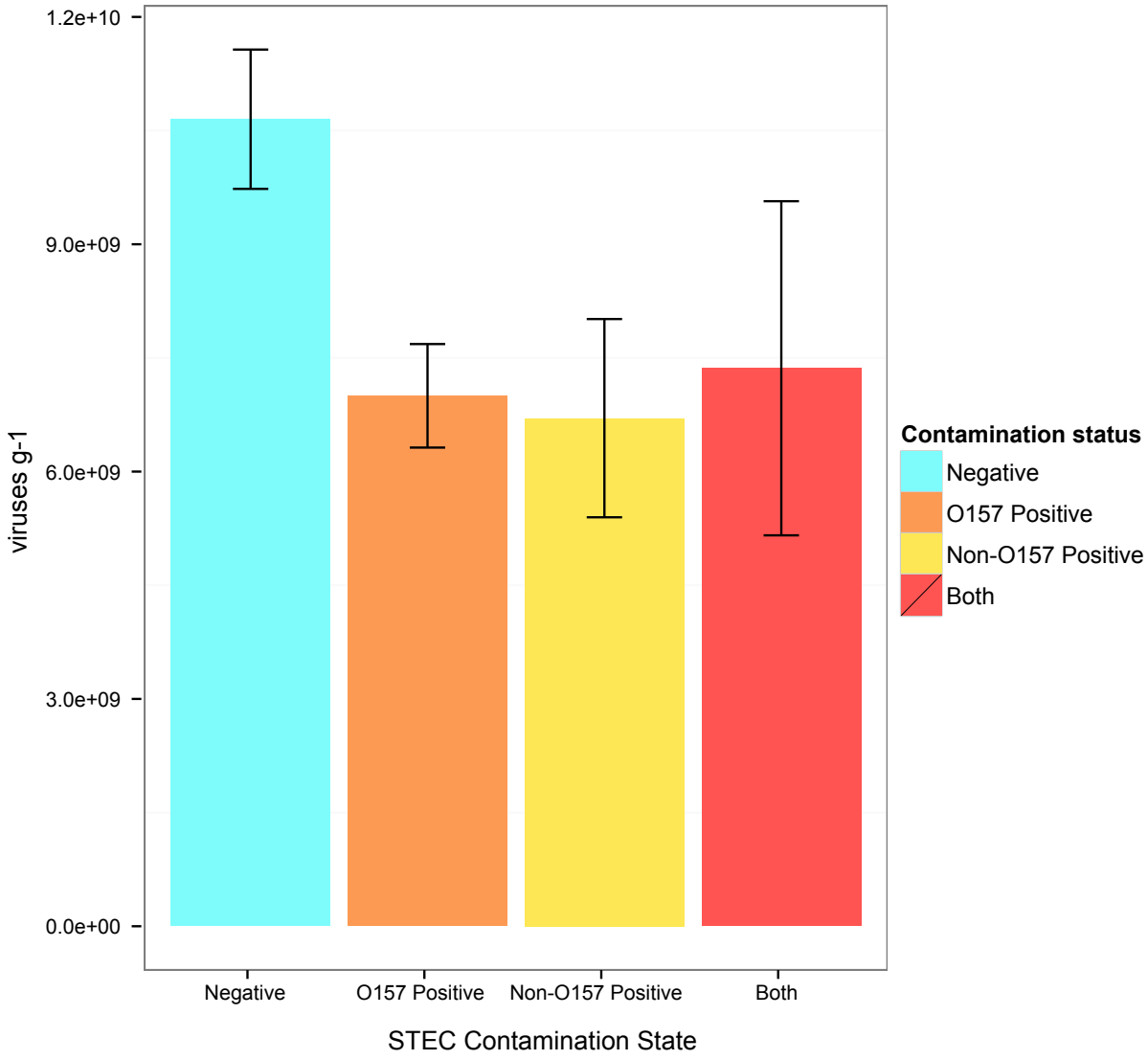


Figure 3.1: Epifluorescence viral counts for fecal viromes. Counts were taken in replicate during processing of fecal mater from pre-harvest cattle at varying states of STEC contamination, which is denoted by bar color: Blue is STEC negative, orange is O157 positive, yellow is non-O157 positive and red is both O157 and non-O157 positive. Error bars represent standard error.

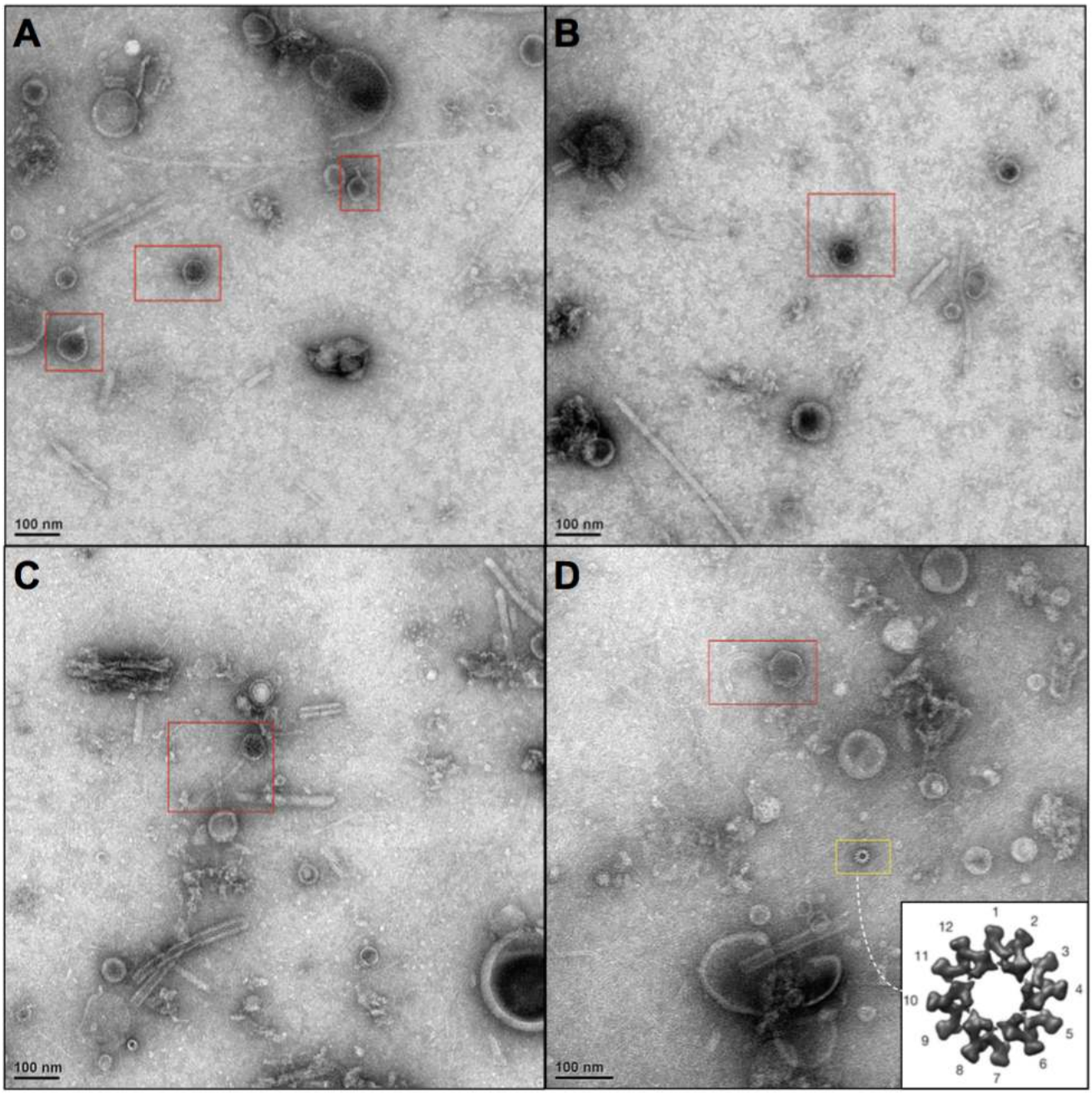


Figure 3.2: Transmission electron microscopy of phages in non-O157 positive fecal sample. Red boxes denote intact phage particles and yellow boxes signify phage parts: (A) Siphoviridae and Podoviridae, (B,C) Siphoviridae, and (D) Siphoviridae and contacted sheath from T4-like Myoviridae as described in Basler et al., 2012.

Table 3.1: Virome sequencing and assembly data

Sample ID	STEC Contamination	Reads	Contigs	ORFs
100	O157 Positive	35,000,438	46,112	114,863
140	O157 Positive	14,142,984	74,128	124,680
117	O157 Positive	25,314,174	111,602	193,934
111	O157 Positive	41,216,564	136,965	188,312
144	Non-O157 Positive	47,960,050	34,858	82,140
141	Non-O157 Positive	2,575,462	18,372	18,280
109	Both	31,426,278	96,945	163,992
118	Both	44,336,940	128,399	207,025
135	Both	26,349,376	254,720	417,854
123	Negative	30,527,942	75,216	126,055
130	Negative	27,474,050	64,022	101,263

14.65%. Finally, the both category samples 109, 118, 135 for Firmicutes and Bacteroidetes were 65.41% and 21.31%, 46.76% and 35.09%, and 69.51% and 17.48%. The ratio of the two phyla varied between individual samples and did not appear to be influenced by STEC metadata category.

At the class level (Fig. 3.3B) the most abundant OTU was Clostridia at greater than 50% of the OTUs in all STEC categories (123 - 64.0%, 130 - 77.80%, 111 - 59.70%, 100 - 68.50%, 140 - 67.40%, 117 - 48.30, 144 - 60.0%, 141 - 58.70%, 109 - 62.0%, 118 - 45.40%, and 135 - 65.40). The second most abundant class was Bacteroidia in all samples, except 123 (negative) in which Spirochaetes had the second largest abundance (12.03% to 14.10%). The other bacterial classes greater than 1.00% were Spirochaetes (except for 130 at .80%) and Bacilli (except for 118 at .60%).

For the fecal viromes, the online web-tool VIROME (Wommack et al., 2012) was used for taxonomic and functional classification (Fig. 3.4 and Table 3.2 and 3.3), however, sample 144 (non-O157) was not available through VIROME and was excluded from this section of the results. The order of dsDNA-tailed phages, Caudvirales, was the majority contributor to viruses in all categories. The other viruses category refers

to viruses outside of Caudovirales, such as ssDNA bacteriophage, eukaryotic viruses (i.e. *Marseillevirus*), and Tectiviridae. The proportion of families within Caudvirales (Siphoviridae, Podoviridae and Myoviridae) also did not appear to be dependent on STEC metadata category. Siphoviridae was the dominant viral family for samples 109 (27.8%; both), 135 (36.7%; both) and 130 (21.0%; negative) (Fig. 3.4A). For these samples Podoviridae was the second most abundant followed by Myoviridae. Podoviridae were the dominant family in samples 123 (21.1%; negative), all the O157 samples (111 - 31.8%, 100 - 36.5%, 140 - 31.5%, and 117 - 37.8%), and 118 (27.4%; both) followed by Siphoviridae and Myoviridae (Fig. 3.4A).

At the genus level, the top 20 phage for each samples were identified (Fig. 3.4B). Within the top 20 *Bacillus* phage, *Clostridium* phage, *Actinomyces* phage, *Lactobacillus* phage, *Croceibacter* phage, *Cellulophaga* phage, *Bacteroides* phage, *Lactococcus* phage, *Streptococcus* phage and *Staphylococcus* phage were present in all the samples, regardless of STEC contamination. Other genera that appeared in the top 20 in at least one of the samples were *Chlamydia* phage (123, 130, 109 and 118), *Microviridae* phage (123, 130, 100, 140, 117, 109, and 118), *Bdellovibrio* phage (123, 130, 117, 109 and 118), *Persicivirga* phage (123, 130, 100, 140, 109 and 118), *Salicola* phage (123, 100 and 140), Dragonfly-associated phage (123 and 130), *Sphingomonas* phage (123, 111, and 118), *Spiroplasma* phage (123 and 130) *Paramecium* phage (123, 130, 111, 140, and 117), *Enterobacteria phage* (111, 100, 140, 117, 109, 118, and 135), *Campylobacter* phage (130, 111 and 118), *Escherichia* phage (111, 117 and 118), *Mycobacterium* phage (111, 100, 117 and 109), *Shigella* phage (111), *Synechococcus* phage (111, 100, 140, 117, 10, and 135), *Pseudomonas* phage (111, 100, 117, 109, 118 and 135), *Tetrasphaera* phage (100 and 140), *Enterococcus* phage (100, 140, 117 and 135), *Listeria* phage (117), *Corynebacteria* phage (135), *Vibrio* phage (130, 109, 118 and 135) and *Aeromonas phage* (135).

Bacillus phage was at the highest relative abundance of identifiable phage for all the samples followed by *Clostridium* phage, except sample 111 (O157 positive) in which *Clostridium* phage was slightly more abundant (18.1% to 17.9%; Fig. 3.4B). In

the remaining samples, *Clostridium* phage was the second most abundant, excluding sample 100 where it was *Cellulophaga* phage (10.2% to 9.6%; Fig. 3.4B). In addition, a large percentage of each sample were unidentified phage.

To determine whether the most abundant phage genera (*Bacillus* and *Clostridium*) were similar to virulent or temperate phage a literature search was conducted on the phage species identified by VIROME (Table 3.2 and 3.3). Those that could not be characterized by the literature were sorted based on 762 positions. In both cases, the majority was similar to virulent phage species. Out of the *Bacillus* group, at least 80% were similar to virulent phage, with an average of 93% virulent among the 9 viromes (Table 3.2). For the *Clostridium*, an average of 71.3% of identified species were virulent between the 9 viromes, with all samples having at least 50% virulent phage (Table 3.3).

3.3.4 Viral functional characteristics

The top annotated functional proteins occurring in all samples were DNA polymerase, single-stranded DNA-binding protein and phage terminase (Table 3.4). In addition, structural proteins (i.e capsid protein and phage tail fiber protein) were among the top annotated proteins in all samples, except 140 and 135 where anaerobic RNR is in the top 5 (O157).

Using VIROME the virulence and pathogenic categories were explored in the SEED database (Overbeek et al., 2005). No STEC associated virulence factors were identified, however, *Staphylococcus aureus* superantigen pathogenicity islands (SaPI) was identified in all samples and *Listeria monocytogenes* pathogenicity island 1 (LIPI-1) was identified in all samples excluding all the STEC negative and O157 positive 111 (Table 3.5). Toxin genes for Streptolysin were identified in samples 135 (both) and 140 (O157). In addition, toxin genes for Streptolysin (140;O157 and 135; both), Chlorea (123; negative) and Diphtheria (135; both) were found.

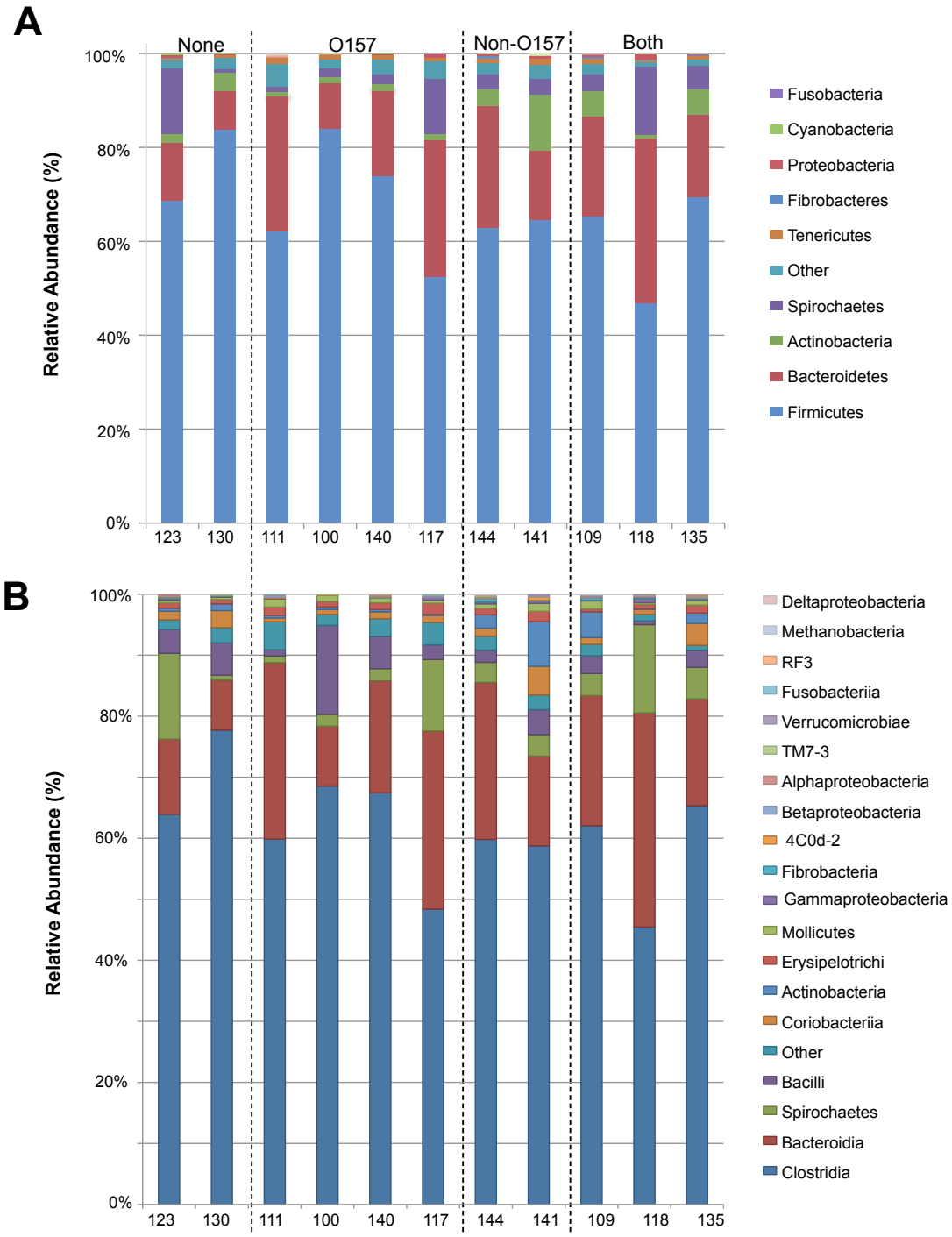


Figure 3.3: Distribution of fecal bacterial taxonomy. Relative abundance charts based on distributions of bacterial phyla (A) and class (B) as a percentage of the total number of classified 16S rRNA sequences. Legend is in order as appears on the bar. Samples are grouped based on STEC infection status and separated by a dotted line (None, Only O157, Non-O157, and Both).

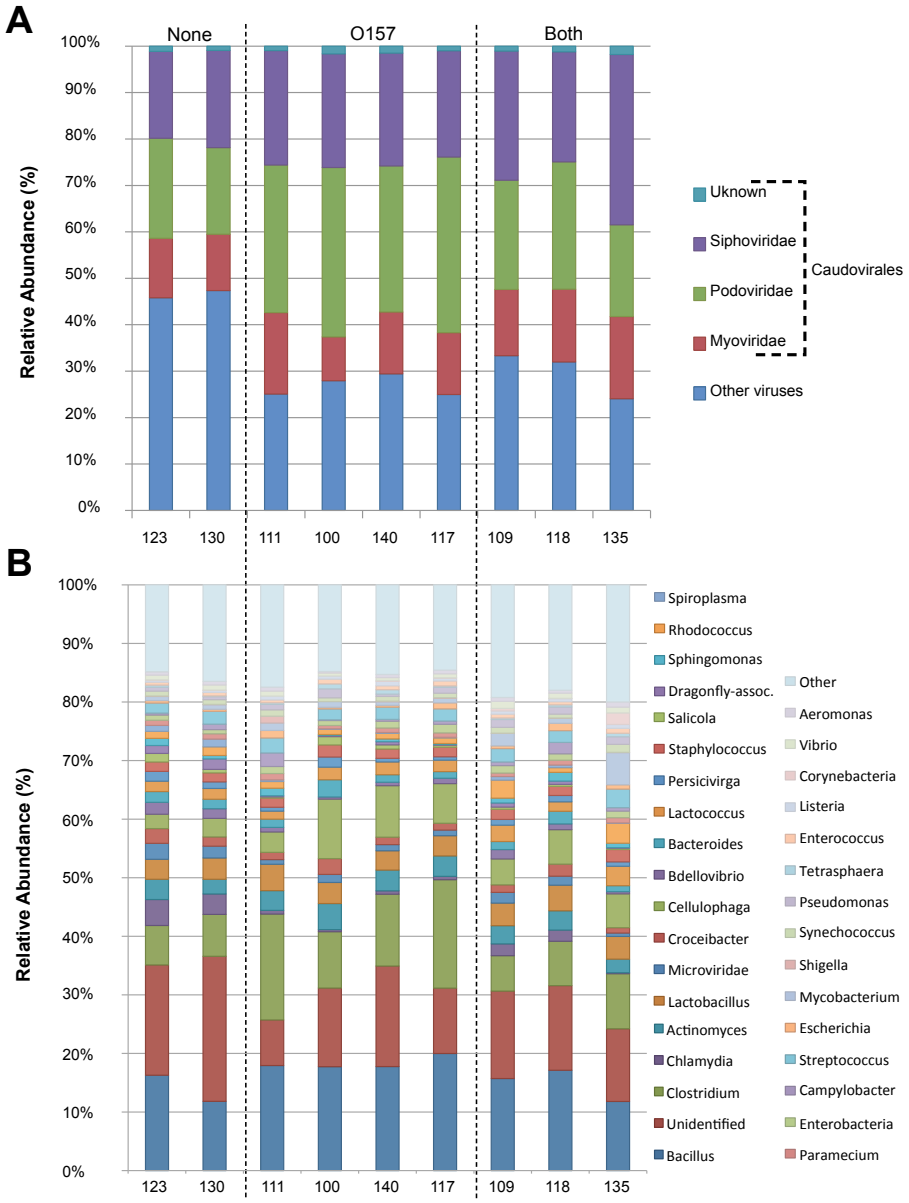


Figure 3.4: Distribution of fecal viral taxonomy. Relative abundance charts as determined by the VIROME web-tool (Wommack et al., 2012) by family (A), whereas (B) compares the top 20 phage genera present in each sample. Samples are grouped based on STEC infection status and separated by a dotted line (None, only O157 and both). The other viruses category refers to viruses outside of Caudovirales. Legend is in order as appears on the bars.

Table 3.2: Lifestyle of homologous *Bacillus* phage species within STEC viromes. Abundance of *Bacillus* species as determine by the VIROME web-tool for each fecal virome. Asterix denotes species characterized by the Polymerase A 762 (*E.coli*) and ND refers to phage whose lifestyles could not be determined. Asterix denotes species characterized by the 762 (*E.coli*) position of their Polymerase A protein. Continued on next page.

Phage Species	**Lifestyle**	**Reference**	**123**	**130**	**111**	**100**	**140**	**117**	**109**	**118**	**135**
SIOphi	ND		13	1	14	1	12	25	16	7	30
virus 1	ND		15	8	13	4	8	16	15	12	27
PM1	ND		2	3	4	1	4	7	3	6	11
Morita2001	ND		0	0	0	0	0	0	0	0	1
AP50	Temperate	Sozhamanna et al., 2008	22	14	15	10	30	38	25	27	31
250	Temperate	Lee et al., 2010	3	8	13	3	7	14	5	13	37
BCJA1c	Temperate	Kropinski et al., 2005	2	3	1	0	3	4	3	3	18
phi105	Temperate	Rutberg et al., 2012	1	1	0	1	2	2	0	1	3
SPbeta	Temperate	Rutberg et al., 2012	1	1	0	0	1	1	1	1	2
IEBH	Temperate	Smeesters et al., 2011	1	1	1	1	0	0	0	0	3
rho11	Temperate	Gnthert et al., 2008	0	0	0	0	0	2	0	0	3
Wip2	Temperate	Schuch et al., 2009	0	0	1	0	0	0	1	0	0
TP21-L	Temperate	Klumpp et al., 2010b	0	0	0	0	1	0	0	0	1
phIS3501	Temperate	Moumen et al., 2012	0	0	0	0	0	1	0	0	9
WBeta	Temperate	Schuch and Fischetti, 2006	0	0	0	0	0	0	0	0	1
SPR	Temperate	Gnthert et al., 2008	0	0	0	0	0	0	0	1	0
SPO2	Temperate	Boice et al., 1969	0	0	0	0	0	0	0	0	1
B103	Virulent	Meijer et al., 2001	217	132	290	121	226	420	302	315	284
phi29	Virulent	Meijer et al., 2001	223	144	315	112	166	414	300	318	263
GA-1	Virulent	Meijer et al., 2001	172	64	173	73	112	226	188	181	187
MG-B1	Virulent	Redondo et al., 2013	113	49	119	49	96	178	138	152	102
G	Virulent	PolA Phe762*	60	51	69	13	33	65	50	48	116
vB_BceM-Bc431v3	Virulent	El-Arab et al., 2013	10	13	27	8	50	52	10	29	98
SPO1	Virulent	Klumpp et al., 2010	33	11	40	8	19	46	50	28	42
M2	Virulent	Meijer et al., 2001	19	14	54	12	20	43	22	33	18
BCD7	Virulent	Gillis et al., 2014	7	7	74	2	0	13	17	91	19
BCP78	Virulent	Lee et al., 2012	145	10	17	7	11	27	10	20	31
SP10	Virulent	PolA Phe762*	12	8	13	1	6	10	31	6	29
phiAGATE	Virulent	PolA Phe762*	4	3	8	2	2	9	12	4	22
0305phi8-36	Virulent	Thomas et al., 2007	3	3	5	3	5	12	4	5	18

Phage Species	Lifestyle	Reference	123	130	111	100	140	117	109	118	135
B4	Virulent	Lee et al., 2013	7	0	4	0	5	5	1	6	18
Bastille	Virulent	Klumpp et al., 2010	7	0	4	3	5	2	6	3	9
W.Ph.	Virulent	Klumpp et al., 2010	2	0	2	3	2	3	3	5	10
SPP1	Virulent	Yasbin et al., 1974	0	1	0	0	1	0	0	1	13
12826	Virulent	Loessner et al., 1997	2	1	1	1	1	2	2	1	4
Curly	Virulent	Lorenz et al., 2013	1	1	4	0	0	5	0	2	0
PBC1	Virulent	Kong et al., 2012	2	1	1	0	0	1	2	0	2
phiE	Virulent	Klumpp et al., 2010	1	1	1	0	0	2	1	2	3
Finn	Virulent	Lorenz et al., 2013	0	0	3	0	0	3	0	1	4
BPS13	Virulent	Shin et al., 2014	0	1	0	0	2	2	0	0	0
Andromeda	Virulent	Lorenz et al., 2013	1	0	1	0	1	2	0	0	0
Eoghan	Virulent	Lorenz et al., 2013	0	0	0	0	0	0	0	1	3
SP82	Virulent	Klumpp et al., 2010	0	0	0	1	0	1	0	0	1
Fah	Virulent	Minakhin et al., 2005	1	0	0	0	0	0	0	1	0
bg2	Virulent	Schmitz et al., 2008	0	0	0	0	0	0	0	0	1
		Total Temperate (ORFs)	30	28	31	15	45	62	35	46	109
		Total Temperate (%)	2.7	5.1	2.4	3.4	5.4	3.8	2.9	3.5	7.4
		Total Virulent (ORFs)	1042	515	1225	419	763	1543	1149	1253	1297
		Total Virulent (%)	94.6	92.8	95.2	95.2	91.7	93.4	94.3	94.6	87.9

Table 3.3: Lifestyle of homologous *Clostridium* phage within STEC viromes. Abundance of *Clostridium* species as determine by the VIROME web-tool for each fecal virome. Asterix denotes species characterized by the Polymerase A 762 (*E.coli*) and ND refers to phage whose lifestyles could not be determined. Asterix denotes species characterized by the 762 (*E.coli*) position of their Polymerase A.

Phage Species	**Lifestyle**	**Reference**	**123**	**130**	**111**	**100**	**140**	**117**	**109**	**118**	**135**
D-1873	Temperate	Sakaguchi et al., 2005	20	28	49	19	57	99	24	55	190
c-st	Temperate	Sakaguchi et al., 2005	33	26	90	17	59	113	24	39	121
phi8074-B1	Temperate	PolA Leu762*	15	30	48	16	55	70	24	51	83
phiCD38-2	Temperate	Fortier and Moineau, 2007	7	10	13	3	26	27	13	5	43
phiCD6356	Temperate	Horgan et al., 2010	2	7	5	7	13	15	4	5	15
phiSM101	Temperate	Nariya et al., 2011	1	1	5	4	5	6	1	4	17
vB_CpeSCP51	Temperate	Gervasi et al., 2013	1	1	4	2	8	6	1	4	7
phi3626	Temperate	Zimmer et al., 2002	3	1	3	1	1	5	2	2	4
phiS63	Temperate	Kim et al., 2012	2	0	4	0	3	8	0	1	6
phiMMP04	Temperate	Meessen-Pinard et al., 2012	2	0	1	0	2	6	0	0	7
phiMMP02	Temperate	Meessen-Pinard et al., 2012	1	0	0	1	4	4	1	1	1
phiCD27	Temperate	Nale et al., 2012	0	0	1	0	0	1	1	0	2
phiC2	Temperate	Fortier and Moineau, 2007	0	1	1	0	0	1	1	0	1
phi24R	Virulent	Morales et al., 2012	75	43	424	29	80	411	56	80	96
CpV1	Virulent	Volozhantsev et al., 2011	68	34	168	26	62	225	68	71	68
phiCTP1	Virulent	Mayer et al., 2010	27	23	127	6	24	84	58	77	140
phiCP7R	Virulent	Volozhantsev et al., 2012	40	26	90	22	31	113	36	38	46
phiZP2	Virulent	Volozhantsev et al., 2012	37	25	111	14	31	105	28	39	34
phiCPV4	Virulent	Volozhantsev et al., 2012	18	18	63	17	30	73	39	27	30
phiCP34O	Virulent	Oakley et al., 2001	10	15	30	17	24	54	30	25	82
phiCP13O	Virulent	Oakley et al., 2001	11	11	18	12	19	30	12	20	60
phiCP26F	Virulent	Seal et al., 2011	8	18	15	13	12	31	18	10	58
phiCP39-O	Virulent	Seal et al., 2011	12	12	13	12	24	27	21	23	33
		Total Temperate (ORFs)	87	105	224	70	233	361	96	167	497
		Total Temperate (%)	22.1	31.8	17.5	29.4	40.9	23.8	20.8	28.9	43.4
		Total Virulent (ORFs)	306	225	1059	168	337	1153	366	410	647
		Total Virulent (%)	77.9	68.2	82.5	70.6	59.1	76.2	79.2	71.1	56.6

Table 3.4: VIROME identified top functional genes. The top five open reading frames (ORFs) with significant homology to a protein with the UniRef 100 plus database with a functional classification as determine by VIROME. Continued on next page.

Sample ID	STEC Contamination	Top Functional Genes	ORFs
100	O157 Positive	Single-stranded DNA-binding protein	269
		Phage terminase, large subunit	255
		Cytosine-specific methyltransferase	209
		DNA polymerase	184
		Phage tail fiber	159
140	O157 Positive	Phage terminase, large subunit	458
		Single-stranded DNA-binding protein	383
		DNA polymerase	344
		Cytosine-specific methyltransferase	317
		Anaerobic ribonucleoside-triphosphate reductase	263
117	O157 Positive	DNA polymerase	721
		Phage terminase, large subunit	602
		Cytosine-specific methyltransferase	559
		Phage tail fiber	526
		Single-stranded DNA-binding protein	494
111	O157 Positive	DNA polymerase	533
		Phage terminase, large subunit	485
		Single-stranded DNA-binding protein	399
		Cytosine-specific methyltransferase	380
		Phage tail fiber	363
109	Both	Phage terminase, large subunit	582
		Single-stranded DNA-binding protein	541
		Phage tail fiber	491
		DNA polymerase	396
		Capsid protein VP1	377

Sample ID	STEC Contamination	Top Functional Genes	ORFs
118	Both	Phage terminase, large subunit	620
		Single-stranded DNA-binding protein	550
		Phage tail fiber	539
		DNA polymerase	510
		Cytosine-specific methyltransferase	485
135	Both	Single-stranded DNA-binding protein	1003
		Anaerobic ribonucleoside-triphosphate reductase	815
		DNA polymerase	663
		Cytosine-specific methyltransferase	562
		Phage terminase, large subunit	136
123	Negative	Single-stranded DNA-binding protein	269
		Phage terminase, large subunit	255
		Cytosine-specific methyltransferase	209
		DNA polymerase	184
		Phage tail fiber	159
130	Negative	Capsid Protein	424
		Phage terminase, large subunit	400
		Single-stranded DNA-binding protein	280
		Structural protein	217
		DNA polymerase	215

Table 3.5: VIROME identified virulence genes. Number within parentheses denotes amount of ORF hits.

Sample ID	**STEC Contamination**	**Pathogencity Islands**	**Toxins**
100	O157 Positive	Listeria LIPI-1 (1) *Staphylococcus* SaPI (5)	
140	O157 Positive	*Listeria* LIPI-1 (1) *Staphylococcus* SaPI (6) (1)	Streptolysin(1)
117	O157 Positive	*Listeria* LIPI-1 (3) *Staphylococcus* SaPI (14)	
111	O157 Positive	*Staphylococcus* SaPI (10)	
109	Both	*Listeria* LIPI-1 (2) *Staphylococcus* SaPI (6)	
118	Both	*Listeria* LIPI-1 (2) *Staphylococcus* SaPI (10)	
135	Both	*Listeria* LIPI-1 (6) *Staphylococcus* SaPI (39)	Streptolysin(2) Diphtheria toxin (1)
123	Negative	*Staphylococcus* SaPI (7)	Chlorea (1)
130	Negative	*Staphylococcus* SaPI (6)	

3.3.5 Marker gene analysis, ribonucleotide reductase

After sorting, RNRs were grouped by class and metadata category (negative, O157, non-O157, both) (Fig. 3.5). The O157 positive samples had an average of 3 Class I RNRs, 22 Class II RNRs, and 112 Class III RNRs. The non-O157 positive sample had 5 Class I RNRs, 10 Class II RNRs, and 53 Class III RNRs. Both O157 and non-O157 positive had an average of 15 Class I RNRs, 33 Class II RNRs, and 158 Class III RNRs. The negative category had no Class I RNRs, an average of 15 Class II RNRs, and 30 Class III RNRs. The counts were normalized by sequencing depth and are depicted in Fig. 3.5.

All the classes of RNR were aligned and clustered separately. At 75% Class I assembled into 11 clusters with no clusters having STEC negative samples (Fig. 3.6) There were 2 clusters containing all categories of STEC positive (O157, non-O157 and both) with similarity to *Enterococcus* phage phiEF24C (cluster #1) and *Pelagibacter* phage HTVC019P (cluster #8). *E.coli* O157 positive had 3 clusters *Listeria* phage spp. (cluster #0,5) and *Cellulophaga* phage phiST (cluster #2). The both only category had 2 clusters *Caulobacter* phage spp. (cluster #3) and *Mycobacterium* phage spp. (cluster #6). Non-O157 and both grouped into 2 clusters both similar to *Staphylococcus* phage SA11 (cluster #4, 9). Non-O157 and O157 positive also grouped into 2 clusters with similarity to *Caulobacter* phage rogue (cluster #7) and *Staphylococcus* phage SA11 (cluster #10). There were no clusters with non-O157 only.

At 75% Class II assembled into 18 clusters. Only 1 cluster had all the categories (non-O157, O157, both and negative), homologous *Roseovarius* sp. 217 phage 1 (cluster #1). In addition, 3 clusters had any of the STEC positive categories and a negative, which were homologous to *Roseovarius* sp. 217 phage 1 (cluster #7), *Dinoroseobacter* phage vBDshPR2C (cluster #2) and *Sulfitobacter* phage phiCB2047-B (cluster #4). STEC negative was found alone in one cluster, *Dickeya* phage Limestone (cluster #17). *E.coli* O157 positive had 3 clusters with similarity to *Clostridium* phage c-st (cluster #16) and *Roseophage* DSS3P2 (cluster #5). The both only category had 6 clusters with homology to *Thermus* phage P2345 (cluster #15), *Mycobacterium* phage Gaia (cluster

#14), and *Sulfitobacter* phage phiCB2047-B (cluster #11). *E.coli* O157 positive and the both category grouped in 2 clusters similar to *Roseophage* spp. (cluster #0) and *Dinoroseobacter* phage vBDshPR2C (cluster #3). *E.coli* non-O157 positive and the both category formed 1 cluster also similar to *Sulfitobacter* phage phiCB2047-B (cluster #9). Only 1 cluster had all three STEC positive categories, homologous to *Sulfitobacter* phage phiCB2047-B (cluster #8).

Class III assembled into 11 clusters at 75%. There were 6 clusters containing all categories (O157, non-O157, both and negative) with similarities to *Clostridium* phage phiCD211 (cluster #2,0), *Aeromonas* phage phiAS5 (cluster #4), *Citrobacter* phage Moogle (cluster #6), *Cronobacter* phage vBCsaPGAP52 (cluster #9), and *Vibrio* phage pVp-1 (cluster #10). In addition, there were 2 clusters containing any of the STEC positive categories and a STEC negative with similarity to *Clostridium* phage phiCD211 (cluster #3) and *Vibrio* phage ICP1 (cluster #7). *E.coli* O157 positive and the both category grouped in 3 clusters with similarity to *Klebsiella* phage KP15 (cluster #5), *Citrobacter* phage Moogle (cluster #8) and *Clostridium* phage phiCD211 (cluster #1). There were no clusters that contained just one of the categories.

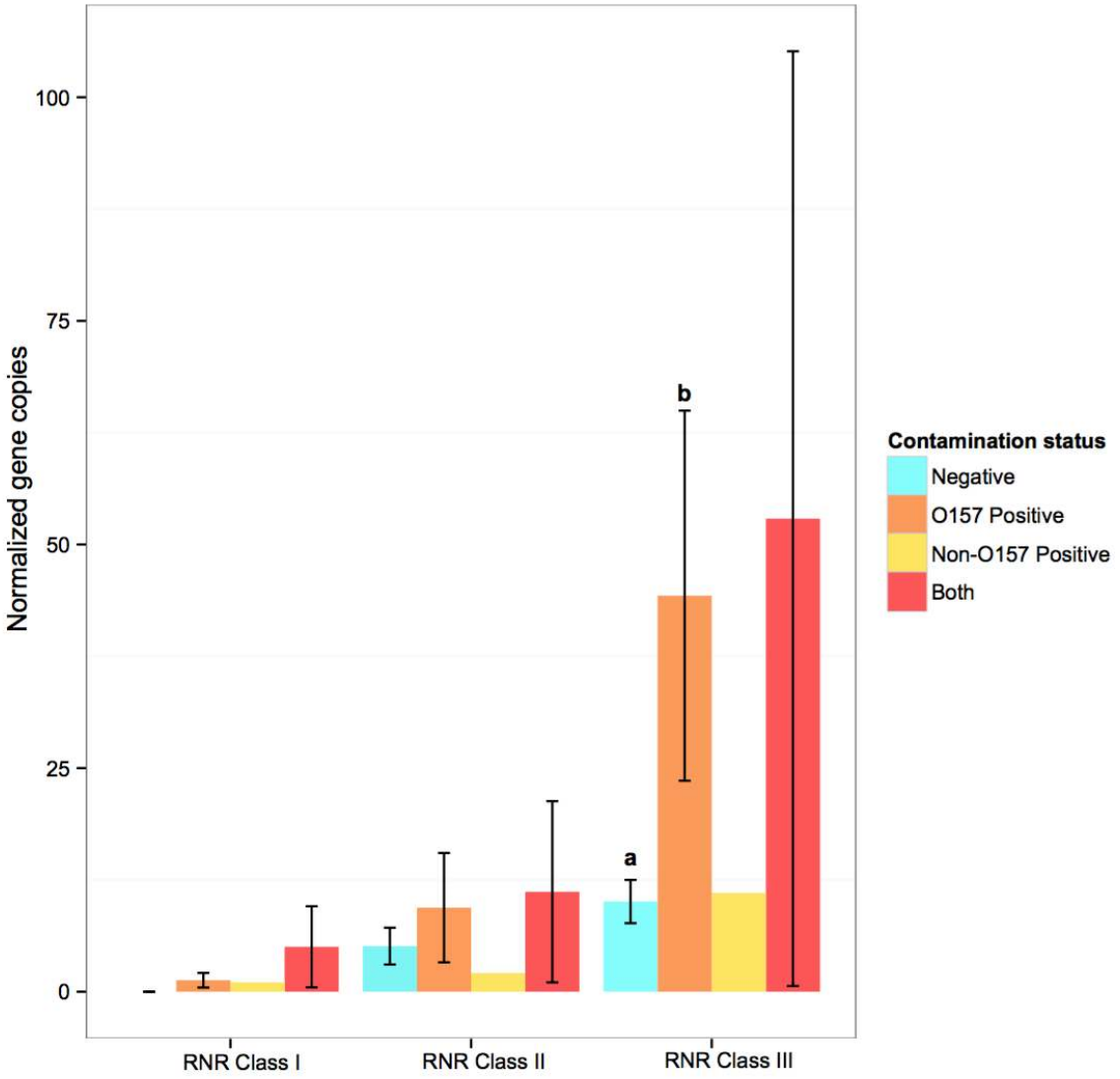


Figure 3.5: Distribution of RNR gene copies by STEC infection status.Viromes grouped by STEC state: O157 positive, non-O157 positive, positive for both and negative for any STEC and RNR class. Header indicates the biochemical class with regards to oxygen dependency. RNRs were normalized by total RNRs/ (total peptides/10,000,000). Error bars represent standard deviation. Bars with different letters were significantly different at an alpha level of 0.05 with Bonferonni correction as determined by the t test. No bar is present in the non-O157 because it was a single sample.

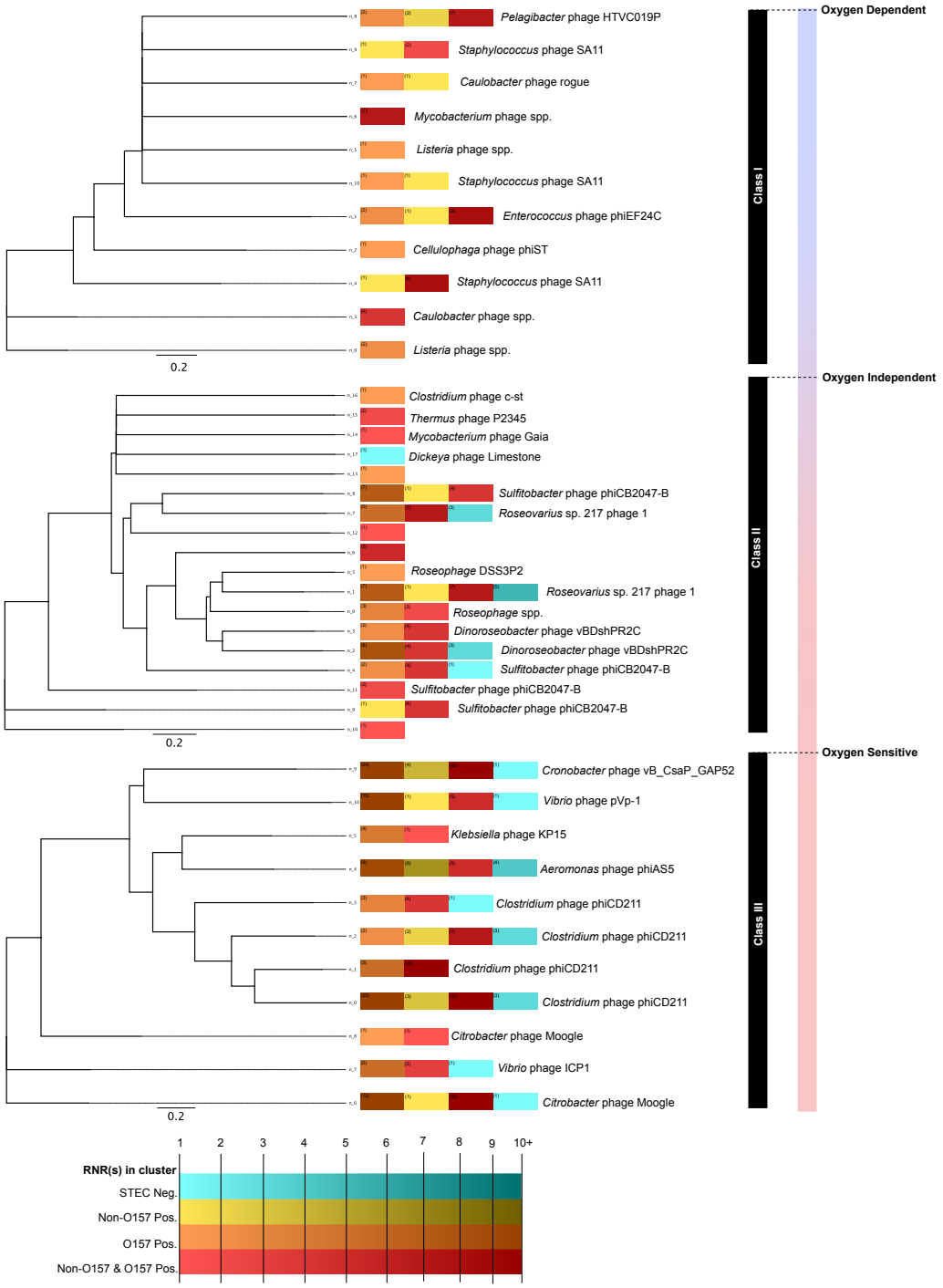


Figure 3.6: Unrooted maximum likelihood trees for Class I, Class II and Class III RNR cluster representatives. Tip boxes represent number of peptides in the cluster (75%), with the darker the shade signifying more peptides within that cluster. Tip nodes are annotated with top BLASTp assignment. The bar to the right of the trees represents oxygen reactivity; Class I RNRs function in aerobic conditions, Class III in anaerobic and Class II at the interface. Scale bars represent amino acid substitutions per site.

3.3.6 Marker gene analysis, polymerase A

PolAs were sorted as in the RNRs, grouped by 762 position and metadata category (Fig. 3.7). The O157 positive samples had an average of 12 Leu762, 16 Phe762, and 1 Tyr762. The non-O157 positive sample had 28 Leu762, 16 Phe762, and no Tyr762. Both O157 and non-O157 positive had an average of 21 Leu762, 22 Phe762, and 2 Tyr762. The negative category had an average of 8 Leu762, 8 Phe762, and no Tyr762. The counts were normalized by sequencing depith and are depicted in Fig. 3.7.

At 75% the PolAs assembled into 50 clusters. There were 4 clusters containing all categories (O157, non-O157, both and negative) with similarities to *Thermus* phage TMA (cluster #13), *Paenibacillus* phage Emery (cluster #33), *Pseudomonas* phage KPP25 (cluster #42) and *Bacillus* phage SP10 (cluster #35). In addition, 4 clusters had any of the STEC positive categories and a negative, which were homologous to *Pseudomonas* phage spp. (cluster #27), *Bacillus* phage spp. (cluster #36), *Mycobacterium* phage spp. (cluster #5), and *Listeria* phage spp. (cluster #10). There was just 1 cluster for negative only (cluster #21), homologous to *Bacillus* phage SP10.

The both only category had 19 clusters with similarity to *Vibrio* phage spp. (clusters #0 and 16), *Mycobacterium* phage spp. (clusters #1, 29 and 30), *Bacillus* phage spp. (clusters #2, 23, 24, 25 and 39), *Listeria* phage LP-026 (cluster #12), *Achromobacter* phage JWF (cluster #26), *Streptococcus* phage P9 (cluster #40), *Clostridium* phage phiCTP1 (cluster #7). *E.coli* O157 positive had 6 clusters *Bacillus* phage spp. (clusters #18 and 45), Endosymbiont phage APSE-1 (cluster #41), *Listeria* phage LP-026 (cluster #11), *Pseudomonas* phage KPP23 (cluster #32), and *Enterobacteria* phage K1F (cluster #20). There was just 1 cluster for non-O157 only (cluster #44) with hits to *Pseudomonas* phage KPP25. In addition, there were 15 clusters with STEC positive (any combination of both, non-O157, and/or O157). These clusters had similarity to *Bacillus* phage spp. (clusters #22, 9, 34 and 38), *Salmonella* phage FSL SP-076 (clusters #17 and 31), *Gordonia* phage GTE2 (cluster #3), *Clostridium* phage spp. (clusters #8 and 14), *Celeribacter* phage P12053L (cluster #19), *Croceibacter*

phage P2559Y (cluster #37), *Pseudomonas* phage KPP25 (cluster #43), *Streptococcus* phage Dp-1 (cluster #6) and *Mycobacterium* phage spp. (cluster #4).

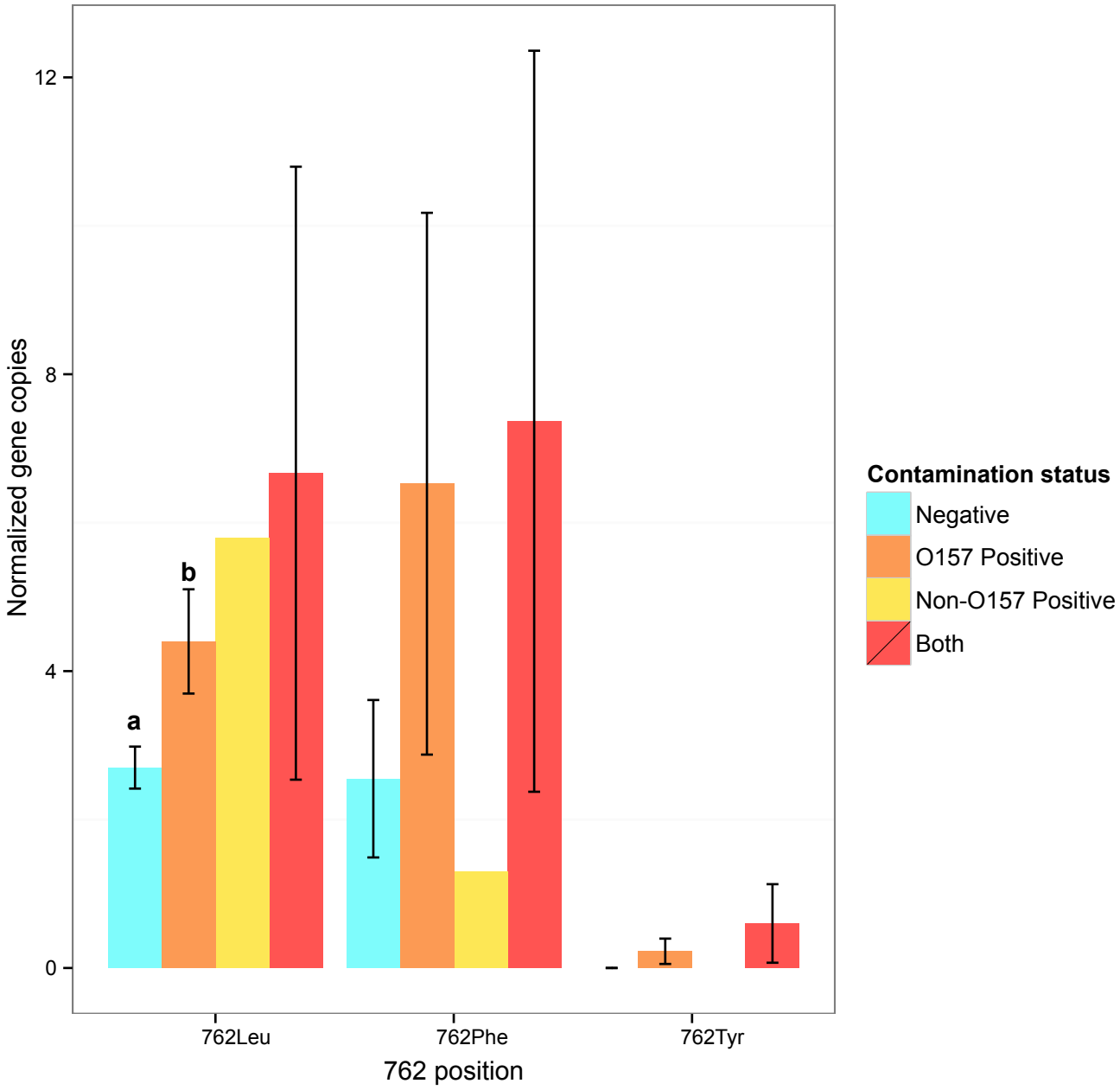


Figure 3.7: Distribution of PolA gene copies by STEC infection status. Bar color denotes varying states of STEC contamination, blue is STEC negative, orange is *E.coli* O157 positive, yellow is non-O157 positive and red is both *E.coli* O157 and non-O157 positive. Bars were grouped by 762 position: Phe762 (lytic), Tyr762 (lytic), and Leu762 (lysogenic). PolAs were normalized by (total PolAs)/ (total peptides/10,000,000).) Error bars represent standard deviation. Bars with different letters were significantly different at an alpha level of 0.05 with Bonferonni correction as determined by the t test. No bar present in the non-O157 because it was a single sample.

Figure 3.8: Unrooted maximum likelihood tree for PolA cluster representatives. Tip boxes represent number of peptides in the cluster, with the darker the shade signifying more peptides within that cluster (75%). Tip nodes are annotated with top BLASTp assignment. Line color represents 762 position (red is Phe762, blue is Tyr762 and purple is Leu762). Scale bar signifies amino acid substitutions per site.

Chapter 4

DISCUSSION

Understanding the complex dyanmics occurring among the viral and bacterial members of a biome may be key in assessing the risk of pathogen invasion. Current research suggests that the indigenous microbial community of the skin is on the forefront of defense against the growth of invading bacteria (Chiller et al., 2001 and Cogen et al., 2008). For instance, *Staphylococcus epidermidis*, a commensal bacterium found on the skin of humans, has demonstrated the ability to prevent the colonization of pathogens, such as *S. aureus* and Group A *Streptococcus*, through the binding of keratinocyte receptors and the production of phenol-soluble modulins and other antimicrobial peptides (Bibel et al., 1983, Cogen et al., 2010a, Cogen et al., 2010b, and Grice et al., 2011). Additionally, increased diversity on healthy intact skin in humans has been speculated to reduce the spread of opportunistic pathogens present in wounds (Gontcharova et al., 2010). This complex and varied ecosystem is poorly understood in cattle and may provide a rich source of information for the mitigation of pathogen invasion and colonization.

Our analyses showcased the relationship between the composition of hide bacterial communities and the incidence of STEC contamination on cattle hides. However, to validate whether these correlations were hide associated and not just a result of varying degrees of fecal contamination, OTUs identified in cattle fecal bacterial communities sampled from the same location were removed to create an artificial hide-only bacterial community. Bacterial families known to inhabit skin and soil environments dominated the hide-only group, while the original dataset also contained dominant families found in the digestion tract of mammals (Fig. 2.2 and 2.3). Alpha diversity metrics indicated an association between low bacterial diversity and STEC occurrence

on cattle hide (Fig. 2.4 and 2.5). Because our data is observational we can only hypothesize about cause and effect. For instance, it is difficult to assess whether a genomic adaption within the pathogen enabled it to outcompete the local hide microbiota or if a disturbance within the microbial community presented STEC with the opportunity to colonize an unoccupied niche.

Shiga-toxigenic *E. coli* is an opportunistic pathogen that carries a number of genes that enable it to survive within the low pH conditions of the stomach and within nutrient-limited environments (van Elsas et al., 2011). Almost 20% of *E. coli* O157 genome is composed of foreign DNA not present in commensal *E. coli* K12 genome (Hayashi T et al., 2001). This additional reservoir of genomic DNA may provide the pathogen with the ability to survive and compete in novel environments, like cattle hide. However, a healthy commensal microbiome may supersede the genomic plasticity of STEC. Generally, antagonism between bacteria can occur indirectly, via competitive exclusion, or directly, by the production of antimicrobial agents or the modification of the environment to unfavorable conditions.

In an established microbial community all the available environmental niches are filled, generating a protective barrier against colonization by an opportunistic pathogen (Hibbing et al., 2010). Therefore, a diverse bacterial community may prevent STEC from taking purchase in a novel environment. Researchers found that competitive interaction with enteric bacteria can displace or inhibit the growth of *E. coli* O157 in vitro (Stavric et al., 1992, Zhao et al., 1998 and 2003). In addition, one study reported the survival of *E. coli* O157 was decreased by 20-30 fold in the presence of *Enterobacter asburiae*, a competitor for the carbon and nitrogen substrates used by the pathogen (Cooley et al., 2006). The reduced diversity that was present in a subset of our hide samples may have enabled STEC to gain a foothold in a vacant niche (Fig. 2.4 and 2.5). Still, commensal hide bacterial species may inhibit STEC through the production of antimicrobial compounds such as organic acids and bacterocins.

In our study, the bacterial OTUs that were at higher relative abundance in the absence of STEC were *Ruminococcus*, *Brevibacterium*, and Nocardioidaceae, while

in the hide-only dataset the bacteria at higher relative abundance were *Soilbacillus* and *Streptococcus* (Fig. 2.8). By looking at the dynamics of the compounds produced by certain species of these bacteria we can begin to decipher the possible impact these groups may have on STEC invasion. Several *Ruminococcus* spp. (e.g. *R. albus*, *R. flavefaciens*) can yield fermentation products succinate acetate, formate and ethanol in the gut, some of which have been shown to impact the survival of STEC (Russel and Rychlik, 2001). *Ruminococcus*, along with *Clostridium* and *Bacterodies*, have also been identified as members of the human microbiota that can inhibit Stx2 production and, thus, potentially limit STEC propagation (deSablet et al., 2009). Interestingly, a species of *Soilbacillus*, *Solibacillus silvestris*, has the ability to degrade N-Acylhomoserine lactones (AHLs) (Morohoshi et al., 2011). AHLs are widely used for the cellular communication phenomena known as quorum sensing in Gram-negative bacteria. *Escherichia coli* do not produce AHLs themselves, but utilize the AHL receptor of the LuxR family, SdiA. SidA recognize AHLs produced from surrounding bacteria to assess the environment and modulate gene expression as necessary (Sperandio, 2010). Previous research has demonstrated that AHL perception in *E. coli* O157 is utilized for intestinal colonization in cattle and is speculated to be critical in *E. coli* survival outside the host (Dziva et al., 2004 and Van Houdt et al., 2006). While the role of AHLs in hide colonization is unknown, *Solibacillus silvestris* and other AHL-degrading bacteria may be interfering with STECs ability to perceive commensal bacteria, thereby decreasing their chances of a successful invasion.

The remaining OTU identified as inversely proportioned between STEC negative and O157 positive is a member of the genus *Streptococcus*. In fact, 7 different *Streptococcus* OTUs were recognized as down regulated in the presence of STEC positive samples in the hide-only dataset (Fig. 2.8). Species of *Streptococcus* (e.g. *Streptococcus thermophiles*, *Streptococcus bovi*, etc.) are lactic acid producers. Lactic acid, in addition to lowering the pH of the environment, has been shown to cause permeabilization of the outer membrane of Gram-negative bacteria, including *E. coli* O157 (Alakomi et al., 2000). Lactic acid producing bacteria (LAB) have been widely

studied as probiotics to inhibit *E.coli* O157, with the majority of these experiments involving *Lactobacillus* spp. (Brashears et al., 2003, Peterson et al., 2007). High levels of lactic acid bacteria were shown to inhibit the growth of *E. coli* O157 in ground beef (Vold et al., 2000) and experimentally infected cattle saw a decrease in *E.coli* O157 shedding when given a probiotic culture of *Streptococcus bovis* and *Lactobacillus gallinarum* (Ohya et al., 2001). Additionally, sheep fed a cocktail of probiotic bacteria including several *Lactobacillus* species, *Streptococcus thermophilus* and *Enterococcus faecium* showed a reduction in non-O157 STEC fecal shedding (Rigobelo et al., 2014). Thus, *Streptococcus* may be a useful in the prevention of STEC colonization in the gut and on the hide. However, more research is needed to assess the complex interactions occurring between the pathogens and the commensal, including phage-host dynamics.

The cattle gut is a complex environment, hosting a diverse population of microbial life (Shanks et al., 2011 and Rice et al., 2012). Despite their role in the evolution of these microbial communities, viruses have been left largely unexplored in STEC research. Therefore, by characterizing the viromes of cattle feces we can hypothesis about the interactions within the microbial network and how it may be facilitating pathogen invasion.

The "kill-the-winner" (ktw) model of phage-host dynamics hypothesizes that the most competitive fast-growing host population will encounter and select for more viruses and, as a result, increase the rate at which the host is killed (Thingstad, 2000). This facilitates turnover of the dominant microbial species, sustaining diversity (Thingstad, 2000). However, not all experimental observations have supported this hypothesis and some reports show abundant viruses infecting rare fast-growing hosts and rare viruses infecting dominant slow-growing hosts (Winter et al., 2010 and Suttle, 2007). To explain this, revisions to the ktw model have been designed to include a "cost of resistance", whereby species gain defense strategies at a cost to fitness (Vage et al., 2013). Because our data is a snapshot of the phage-host dynamics it is difficult to fit into a model. However, we can speculate that because the phage at the highest abundance in the viromes (*Bacillus* phage and *Clostridium* phage; Fig. 3.4)

had hosts within the same class at high relative abundance in the matching 16S rRNA libraries (Clostridia and Bacilli; Fig. 3.3) that there is active phage lysis of some of these hosts. This is even more convincing due to the high percentage of virulent phage species within the *Bacillus* and *Clostridium* phage groups (Table 3.2 and 3.3). The high turnover in the dominant phage species may provide the conditions for a pathogen or other commensal species to takeover a newly vacant niche. Therefore, it is important to look at other members of the phage population that may be acting on competing bacterial species.

Ribonucleotide reductases (RNRs) are found chiefly in virulent phage and the class of RNR found within viromes reflects the oxygen availability and host bacteria present within the ecosystem sampled (Dwivedi et al., 2013). Class III RNRs are typically absent in marine viromes, but have been identified in environments such as hot springs, fermented foods, low/medium salinity solar salterns, and freshwater metagenomes (Dwivedi et al., 2013). In our viromes, the Class III RNRs were the most abundant class of RNR found, regardless of STEC contamination state (Fig. 3.5). They were even among the top 5 functional proteins found in O157 positive sample 140 and O157 and non-O157 positive samples 135 (Table 3.4). The Class III RNR clusters were largely shared among all the STEC categories, with similarity to phage infecting facultative anaerobes (*Cronobacter*, *Vibrio*, *Aeromonas*, and *Citrobacter*) and anaerobe (*Clostridium*) (Fig. 3.6). This reflects the anaerobic nature of the gut, whereby phage with only class III RNR genes infect strict anaerobes, while phage with class I and III RNRs generally can infect facultative anaerobes (Guarner and Malagelada, 2003 and Dwivedi et al., 2013).

Class I and II RNRs primarily come from phage infecting aerobes or facultative anaerobes and have been identified in various seawater metagenomes (Dwivedi et al., 2013). We found very few Class I RNRs, with none identified in STEC negative samples (Fig. 3.5). One of the clusters of the Class I RNRs had homology to *Enterococcus* phage phiEF24C, which has been shown to be virulent against *Enterococcus faecalis*, a commensal bacterium found in the gastrointestinal tract of mammals (Fig. 3.6). It

has been revealed to have a broad host range, a short latent stage and a large burst size (Uchiyama et al., 2008), making them suitable to diminish commensals in the gut and provide a means for STEC invasion. In addition, three non-O157 containing clusters for of Class I RNRs have similarity to *Staphylococcus* phage SA11. This agrees with the reduction in relative abundance of *Staphylococcus* OTUs in non-O157 positive hide samples (Fig. 2.8). Thus, we are begining to see possible phage-host interactions between the fecal virome and the hide 16S communities.

In the fecal viromes, DNA polymerase was among the top 5 functional proteins in all libraries (Table 3.4). Due to their biochemistry, we can infer information about phage lifestyle from looking at the PolA 762 position. For instance, the fecal viromes have a large distribution of viral PolAs with a Phe762 position or a Phe762Leu position. The Phe762 are generally lytic phage, characteristic of *Enterobacteria* phage T5, *Bacillus* phage SP10/ SP01, *Lactobacillus* phage Lb338-1, whereas the Phe762Leu position are typical of lysogenic phage, like *Roseobacter* phage RDJL Phi 1 and *Edwardsiella* phage eiDWF. The Phe762Tyr PolAs are limited within our fecal viromes, only identified in *E. coli* O157 positive and the both category. These are usually highly virulent and can be found in the genomes of phage such as *Enterobacteria* phage T3 and T7 and *Cyanophage* S-CBP3 (Schmidt et al., 2014).

The PolA clusters were similar to phage found in the environment and in the gut (Fig. 3.8). A majority of the PolA clusters have similarity to *Bacillus* phage spp., regardless of STEC contamination state and largely among the Phe762 clusters (Fig. 3.8). *Bacillus* phage are not uncommon in humans, animals and the environment and were determined to be the largest identified taxonomic group of phage in our viromes (Breibart et al., 2008, Schuch et al., 2010, Haynes and Rohwer, 2011; Fig. 3.7 and 3.8). A notable phage species similar to one of the few virulent Phe762Tyr STEC positive clusters was *Enterobacteria* phage K1F (Fig. 3.8), a member of the genera of phage (*Enterobacteria*) that were among the top 20 abundant phage for all STEC positive categories (Fig. 3.8. *Enterobacteria* phage K1F requires the K1 polysaccharide capsule for infection. This K1 antigen is found in both generic and pathogenic *E.coli*, but is not

carried by STEC (Scholl and Merril 2005). The appearance of this phage may indicate the lysis of competitive *E.coli* with the K1 antigen. Furthermore, two clusters of PolA, which were only found in the lysogenic STEC positive fecal samples, were similar to *Streptococcus* phage Dp-1 and P9 (Fig. 3.8). *Streptococcus* phage Dp-1 is known to produce an amidase phage-associated-lysin active against at least 15 serotypes of *Streptococcus pneumococci* and evolutionarily similar to the lysin found in *Lactococcous lactis* infecting phage (Sheehan et al., 1997, Loeffler et al. 2001 and O'Flaherty et al., 2009). This powerful enzyme, in combination with the other *Streptococcus* phage species found in STEC positive samples, may explain the reduction in *Streptococcus* OTUs found in STEC positive samples (Fig. 2.8). This also agrees with the overall gene copy counts, in which the Phe762Leu phage in STEC negative samples were significantly less abundant than the O157 positive Phe762Leu phage (Fig. 3.7).

Phage are also responsible for the exchange of genetic material, driving the evolution of new pathogens. In fact, lambdoid bacteriophages carrying the stx gene (Stx phage) are recognized to have conferred pathogenicity to *E. coli* O157:H7 (O'Brien et al., 1983). Stx phage incorporate the Shiga-like toxin producing genes (stx1 and stx2) into the host genome through the lysogenic replication strategy (Allison, 2007) and release the toxin upon entry into the lytic cycle (Wagner et al 2002 and Bergan et al., 2012). Previous research has determined that the presence of stx-phage can influence the indigenous bacterial community by infecting commensal gut *E. coli* populations, thus propagating the diseased state (Gamage et al., 2003). This is especially significant because several stx-phage are known to have a broad host range that enable it to infect *E. coli* spp. and potentially *Enterococcus* and *Shigella* spp as well (James et al., 2001). We were unable to detect any STEC virulence genes within our viromes, most likely due to their rarity within populations, lysogenic lifestyle and lack of database completeness (Allison et al., 2007). However, other genes were found that are associated with pathogenic bacteria. For instance, superantigen-encoding pathogenicity islands of *Staphylococcus aureus* (SaPIs) were detected in all the viromes (Table 3.5). These superantigens induce disease in humans (i.e toxic shock syndrome), as well as cattle

(i.e. mastisis) (Lindsay et al., 1998 and Fueyo et al., 2005). In addition, all the STEC O157 positive samples, except 111, had ORFs with similarity to the *Listeria* pathogenicity island, LIPI-1, present in *L. monocytogenes* and *L. ivanovii* (Table 3.5). LIPI-1 encodes the hemolysin Listeriolysin O (LLO) responsible for Listeriosis (Cossart, 1988 and Vzquez-Boland et al., 2001). These data suggests that the exploration of viral communities may enlighten not only our knowledge of STEC and its prevalence on cattle hide, but also other human and animal pathogens that may be lurking in your hamburger.

Overall, these data indicates that the microbial community diversity and viral interaction is actively linked to the survival of STEC on the hide. Additional research is needed to assess the validity of some of the hypotheses presented in this thesis. For instance, a bacterial metagenomic survey and enzyme assay may provide evidence for the contribution of AHL-degrading bacteria in the decline of STEC on the hide. With further elucidation we could use the information from this research to build risk assessment tools from community profiles. Individual OTUs could be used as biotherapeutic agents to mitigate STEC contamination and targeted phage knockout could be used to bolster a strong diverse community of STEC-limiting commensals, thereby providing a natural method for pathogen control and food protection.

REFERENCES

Alakomi, H-L., E. Skytt, M. Saarela, T. Mattila-Sandholm, K. Latva-Kala, and I. M. Helander. 2000. Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. *Applied and environmental microbiology* 66: 2001-2005.

Allison, Heather E. 2007. Stx-phages: drivers and mediators of the evolution of STEC and STEC-like pathogens.*Future Microbiology* 2: 165-174.

Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman. 1990. Basic local alignment search tool. *Journal of molecular biology* 215: 403-410.

Anderson, Rika E., Mitchell L. Sogin, and John A. Baross. 2014. Evolutionary strategies of viruses, bacteria and archaea in hydrothermal vent ecosystems revealed through metagenomics. *PloS one* 9:e109696.

Angly, Florent E., Ben Felts, Mya Breitbart, Peter Salamon, Robert A. Edwards, Craig Carlson, Amy M. Chan et al. 2006. The marine viromes of four oceanic regions. *PLoS biol* 4: e368.

Arthur, Terrance M., Joseph M. Bosilevac, Dayna M. Brichta-Harhay, Michael N. Guerini, Norasak Kalchayanand, Steven D. Shackelford, Tommy L. Wheeler, and Mohammad Koohmaraie. 2007. Transportation and lairage environment effects on prevalence, numbers, and diversity of *Escherichia coli* O157: H7 on hides and carcasses of beef cattle at processing. *Journal of Food Protection* 70: 280-286.

Bai, Jianfa, Xiaorong Shi, and T. G. Nagaraja. 2010. A multiplex PCR procedure for the detection of six major virulence genes in *Escherichia coli* O157: H7. *Journal of microbiological methods* 82: 85-89.

Bai, Jianfa, Zachary D. Paddock, Xiaorong Shi, Shubo Li, Baoyan An, and Tiruvoor G. Nagaraja. 2012. Applicability of a multiplex PCR to detect the seven major Shiga toxin producing *Escherichia coli* based on genes that code for serogroup-specific O-antigens and major virulence factors in cattle feces. *Foodborne pathogens and disease* 9: 541-548.

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455-477.

Barkocy-Gallagher, Genevieve A., Terrance M. Arthur, Mildred Rivera-Betancourt, Xiangwu Nou, Steven D. Shackelford, Tommy L. Wheeler, and Mohammad Koohmaraie. 2003. Seasonal prevalence of Shiga toxin-producing *Escherichia coli*, including O157: H7 and non-O157 serotypes, and *Salmonella* in commercial beef processing plants. *Journal of Food Protection* 66: 1978-1986.

Basler, M, M. Pilhofer, G. P. Henderson, G. J. Jensen, and J. J. Mekalanos. 2012. Type VI secretion requires a dynamic contractile phage tail-like structure. *Nature* 483: 182-186.

Bergan, Jonas, Anne Berit Dyve Lingelem, Roger Simm, Tore Skotland, and Kirsten Sandvig. 2012. Shiga toxins. *Toxicon* 60: 1085-1107.

Bell R. 1997. Distribution and sources of microbial contamination on beef carcasses. *Journal of Applied Microbiology* 82:292-300.

Bench, Shellie R., Thomas E. Hanson, Kurt E. Williamson, Dhritiman Ghosh, Mark Radosovich, Kui Wang, and K. Eric Wommack. 2007. Metagenomic characterization of Chesapeake Bay virioplankton. *Applied and environmental microbiology* 73: 7629-7641.

Berg Miller, Margret E., Carl J. Yeoman, Nicholas Chia, Susannah G. Tringe, Florent E. Angly, Robert A. Edwards, Harry J. Flint, Raphael Lamed, Edward A. Bayer, and Bryan A. White. 2012. Phage bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome. *Environmental microbiology* 14: 207-227.

Bibel, Debra Jan, Raza Aly, Charlene Bayles, Walter G. Strauss, Henry R. Shinefield, and Howard I. Maibach. 1983. Competitive adherence as a mechanism of bacterial interference. *Canadian journal of microbiology* 29: 700-703.

Boice, Lu Belle. 1969. Evidence that *Bacillus subtilis* Bacteriophage SPO2 is Temperate and Heteroimmune to Bacteriophage phi105. *Journal of virology* 4: 47-49.

Boyce TG, Swerdlow DL, Griffin PM. 1995. *Escherichia coli* O157: H7 and the hemolytic uremic syndrome. *New England Journal of Medicine* 333: 364-368.

Brashears, M. M., D. Jaroni, and J. Trimble. 2003. Isolation, selection, and characterization of lactic acid bacteria for a competitive exclusion product to reduce shedding of *Escherichia coli* O157: H7 in cattle. *Journal of Food Protection* 66: 355-363.

Breitbart, Mya, Peter Salamon, Bjarne Andresen, Joseph M. Mahaffy, Anca M. Segall, David Mead, Farooq Azam, and Forest Rohwer. 2002. Genomic analysis of uncultured marine viral communities. *Proceedings of the National Academy of Sciences* 99: 14250-14255.

Breitbart, Mya, Ian Hewson, Ben Felts, Joseph M. Mahaffy, James Nulton, Peter Salamon, and Forest Rohwer. 2003. Metagenomic analyses of an uncultured viral community from human feces. *Journal of bacteriology* 185: 6220-6223.

Breitbart, Mya, Ben Felts, Scott Kelley, Joseph M. Mahaffy, James Nulton, Peter Salamon, and Forest Rohwer. 2004a. Diversity and population structure of a nearshore marine sediment viral community. *Proceedings of the Royal Society of London B: Biological Sciences* 271: 565-574.

Breitbart, Mya, Jon H. Miyake, and Forest Rohwer. 2004b. Global distribution of nearly identical phage-encoded DNA sequences. *FEMS microbiology letters* 236: 249-256.

Breitbart, Mya, Matthew Haynes, Scott Kelley, Florent Angly, Robert A. Edwards, Ben Felts, Joseph M. Mahaffy et al. 2008. Viral diversity and dynamics in an infant gut. *Research in microbiology* 159: 367-373.

Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM, Strockbine NA. 2005. Non-O157 Shiga toxin producing *Escherichia coli* infections in the United States, 1983-2002. *Journal of Infectious Diseases* 192: 1422-1429.

Buffie CG, Pamer EG. 2013. Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology* 13: 790-801.

Callaway, Todd R., M. A. Carr, T. S. Edrington, Robin C. Anderson, and David J. Nisbet. 2009. Diet, *Escherichia coli* O157: H7, and cattle: a review after 10 years. *Current issues in molecular biology* 11: 67.

Cann, Alan James, Sarah Elizabeth Fandrich, and Shaun Heaphy. 2005. Analysis of the virus population present in equine faeces indicates the presence of hundreds of uncharacterized virus genomes. *Virus genes* 30: 151-156.

Caporaso, J. Gregory, Justin Kuczynski, Jesse Stombaugh, Kyle Bittinger, Frederic D. Bushman, Elizabeth K. Costello, Noah Fierer et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* 7: 335-336.

Caprioli A, Morabito S, Brugre H, Oswald E. 2005. Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Veterinary research* 36: 289-311.

Chao, Anne. 1984. Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*265-270.

Chiller, Katarina, Bryan A. Selkin, and George J. Murakawa. 2001. Skin microflora and bacterial infections of the skin. *In Journal of Investigative Dermatology Symposium Proceedings* 6: 170-174

Clokie, Martha RJ, Andrew D. Millard, Andrey V. Letarov, and Shaun Heaphy. 2011. Phages in nature. *Bacteriophage*1: 31-45.

Cogen, A. L., V. Nizet, and R. L. Gallo. 2008. Skin microbiota: a source of disease or defence? *British Journal of Dermatology* 158: 442-455.

Cogen, Anna L., Kenshi Yamasaki, Katheryn M. Sanchez, Robert A. Dorschner, Yuping Lai, Daniel T. MacLeod, Justin W. Torpey et al 2010a. Selective antimicrobial action is provided by phenol-soluble modulins derived from *Staphylococcus epidermidis*, a normal resident of the skin. *Journal of Investigative Dermatology* 130: 192-200.

Cogen, Anna L., Kenshi Yamasaki, Jun Muto, Katheryn M. Sanchez, Laura Crotty Alexander, Jackelyn Tanios, Yuping Lai, Judy E. Kim, Victor Nizet, and Richard L. Gallo. 2010b. *Staphylococcus epidermidis* antimicrobial delta-toxin (phenol-soluble modulin-gamma) cooperates with host antimicrobial peptides to kill group A *Streptococcus*. *PloS one* 5: e8557-e8557.

Collis, V. J., C-A. Reid, M. L. Hutchison, M. H. Davies, K. P. A. Wheeler, A. Small, and Ands Buncic. 2004. Spread of marker bacteria from the hides of cattle in a simulated livestock market and at an abattoir. *Journal of Food Protection* 67: 2397-2402.

Cooley, Michael B., Diana Chao, and Robert E. Mandrell. 2006. *Escherichia coli* O157: H7 survival and growth on lettuce is altered by the presence of epiphytic bacteria. *Journal of Food Protection* 69: 2329-2335.

Cossart, P. 1988. The listeriolysin O gene: a chromosomal locus crucial for the virulence of *Listeria monocytogenes*. *Infection* 16: S157-S159.

de Sablet, Thibaut, Christophe Chassard, Annick Bernalier-Donadille, Marjolaine Vareille, Alain P. Gobert, and Christine Martin. 2009. Human microbiota-secreted factors inhibit shiga toxin synthesis by enterohemorrhagic *Escherichia coli* O157: H7. *Infection and immunity* 77: 783-790.

DeSantis, Todd Z., Philip Hugenholtz, Neils Larsen, Mark Rojas, Eoin L. Brodie, Keith Keller, Thomas Huber, Daniel Dalevi, Ping Hu, and Gary L. Andersen. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology* 72: 5069-5072.

Dewsbury, Diana Marisa Adele. 2015. Epidemiology of Shiga Toxin-producing *Escherichia coli* in the Bovine Reservoir: Seasonal Prevalence and Geographic Distribution. PhD diss., Kansas State University.

Doublie, Sylvie, Stanley Tabor, Alexander M. Long, Charles C. Richardson, and Tom Ellenberger. 1998. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 resolution. *Nature* 391: 251-258.

Dwivedi, Bhakti, Bingjie Xue, Daniel Lundin, Robert A. Edwards, and Mya Breitbart. 2013. A bioinformatic analysis of ribonucleotide reductase genes in phage genomes and metagenomes. *BMC evolutionary biology* 13: 33.

Dziva, Francis, Pauline M. van Diemen, Mark P. Stevens, Amanda J. Smith, and Timothy S. Wallis. 2004. Identification of *Escherichia coli* O157: H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. *Microbiology* 150: 3631-3645.

Edgar, Robert C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460-2461.

El-Arabi, Tarek F., Mansel W. Griffiths, Yi-Min She, Andre Villegas, Erika J. Lingohr, and Andrew M. Kropinski. 2013. Genome sequence and analysis of a broad-host range lytic bacteriophage that infects the *Bacillus cereus* group. *Virol J* 10: 10-48.

Elder, Robert O., James E. Keen, Gregory R. Siragusa, Genevieve A. Barkocy-Gallagher, Mohammad Koohmaraie, and William W. Laegreid. 2000. Correlation of enterohemorrhagic *Escherichia coli* O157 prevalence in feces, hides, and carcasses of beef cattle during processing. *Proceedings of the National Academy of Sciences* 97: 2999-3003.

Elliott, E. J., R. M. Robins-Browne, E. V. O'loughlin, V. Bennett-Wood, J. Bourke, P. Henning, G. G. Hogg, J. Knight, H. Powell, and D. Redmond. 2001. Nationwide study of haemolytic uraemic syndrome: clinical, microbiological, and epidemiological features. *Archives of disease in childhood* 85: 125-131.

Fadrosh, Douglas W., Bing Ma, Pawel Gajer, Naomi Sengamalay, Sandra Ott, Rebecca M. Brotman, and Jacques Ravel. 2014. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* 2: 1-7.

Faith, Daniel P. 1992. Conservation evaluation and phylogenetic diversity. *Biological conservation* 61: 1-10.

Fierer, Noah, Mya Breitbart, James Nulton, Peter Salamon, Catherine Lozupone, Ryan Jones, Michael Robeson et al. 2007. Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of bacteria, archaea, fungi, and viruses in soil. *Applied and environmental microbiology* 73: 7059-7066.

File, Jonathan, Franoise Ttart, Curtis A. Suttle, and H. M. Krisch. 2005. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proceedings of the National Academy of Sciences of the United States of America* 102: 12471-12476.

Fortier, Louis-Charles, and Sylvain Moineau. 2007. Morphological and genetic diversity of temperate phages in *Clostridium difficile*. *Applied and environmental* microbiology 73: 7358-7366.

Fueyo, J. M., M. Carmen Mendoza, M. Rosario Rodicio, J. Muniz, M. A. Alvarez, and M. Cruz Martn. 2005. Cytotoxin and pyrogenic toxin superantigen gene profiles of *Staphylococcus aureus* associated with subclinical mastitis in dairy cows and relationships with macrorestriction genomic profiles. *Journal of clinical microbiology* 43: 1278-1284.

Gamage, Shantini D., Jane E. Strasser, Claudia L. Chalk, and Alison A. Weiss. 2003. Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infection and immunity* 71: 3107-3115.

Gerber, Angela, Helge Karch, Franz Allerberger, Hege M. Verweyen, and Lothar B. Zimmerhackl. 2002. Clinical course and the role of Shiga toxin producing *Escherichia coli* infection in the hemolytic-uremic syndrome in pediatric patients, 1997 - 2000, in Germany and Austria: a prospective study. *Journal of Infectious Diseases* 186: 493-500.

Gervasi, Teresa, Rosario Lo Curto, Arjan Narbad, and Melinda J. Mayer. 2013. Complete genome sequence of phiCP51, a temperate bacteriophage of *Clostridium perfringens*. *Archives of virology* 158: 2015-2017.

Gillis, Annika, and Jacques Mahillon. 2014. Phages preying on *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: past, present and future. *Viruses* 6: 2623-2672.

Gontcharova, Viktoria, Eunseog Youn, Yan Sun, Randall D. Wolcott, and Scot E. Dowd. 2010. A comparison of bacterial composition in diabetic ulcers and contralateral intact skin. *The open microbiology journal* 4: 8.

Grice, Elizabeth A., and Julia A. Segre. 2011. The skin microbiome. *Nature Reviews Microbiology* 9: 244-253.

Guarner, Francisco, and Malagelada, Juan-R. 2003. Gut flora in health and disease. *The Lancet*: 512-519.

Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic biology* 52: 696-704.

Gnthert, Ursula, Luzia Reiners, and Roland Lauster. 1986. Cloning and expression of *Bacillus subtilis* phage DNA methyltransferase genes in *Escherichia coli* and *B. subtilis*. *Gene* 41: 261-270.

Hancock, D. D., T. E. Besser, M. L. Kinsel, P. I. Tarr, D. H. Rice, and M. G. Paros. 1994. The prevalence of *Escherichia coli* O157. H7 in dairy and beef cattle in Washington State. *Epidemiology and Infection* 113: 199-207.

Hayashi, Tetsuya, Kozo Makino, Makoto Ohnishi, Ken Kurokawa, Kazuo Ishii, Katsushi Yokoyama, Chang-Gyun Han et al. 2001. Complete genome sequence of enterohemorrhagic *Eschelichia coli* O157: H7 and genomic comparison with a laboratory strain K-12. *DNA research* 8: 11-22.

Haynes, Matthew, and Forest Rohwer. 2011. The human virome. In Metagenomics of the human body. *Springer New York* pp. 63-77.

Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010. Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews Microbiology* 8: 15-25.

Horgan M, O'Sullivan O, Coffey A, Fitzgerald GF, van Sinderen D, McAuliffe O, Ross RP. 2012. Genome analysis of the *Clostridium difficile* phage PhiCD6356, a temperate phage of the Siphoviridae family. *Gene* 462: 34-43.

Hughes JM, Wilson ME, Johnson KE, Thorpe CM, Sears CL. 2006. The emerging clinical importance of non-O157 Shiga toxin producing *Escherichia coli*. *Clinical Infectious Diseases* 43: 1587-1595.

Hussein, H. S. 2007. Prevalence and pathogenicity of Shiga toxin-producing in beef cattle and their products. *Journal of Animal Science* 85: e63-e72.

James, Chloe E., Karen N. Stanley, Heather E. Allison, Harry J. Flint, Colin S. Stewart, Richard J. Sharp, Jon R. Saunders, and Alan J. McCarthy. 2001. Lytic and Lysogenic Infection of Diverse *Escherichia coli* and *Shigella* Strains with a Verocytotoxigenic Bacteriophage. 2001. *Applied and environmental microbiology* 67: 4335-4337.

Jami, Elie, and Itzhak Mizrahi. 2012. Composition and similarity of bovine rumen microbiota across individual animals. *PloS one* 7: e33306.

Jami, Elie, Bryan A. White, and Itzhak Mizrahi. 2014. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PloS one* 9: e85423

Jiang, Xiuping, Jennie Morgan, and Michael P. Doyle. 2002. Fate of *Escherichia coli* O157: H7 in manure-amended soil. *Applied and environmental microbiology* 68: 2605-2609.

John, Seth G., Carolina B. Mendez, Li Deng, Bonnie Poulos, Anne Kathryn M. Kauffman, Suzanne Kern, Jennifer Brum, Martin F. Polz, Edward A. Boyle, and Matthew B. Sullivan. 2011. A simple and efficient method for concentration of ocean viruses by chemical flocculation. *Environmental microbiology reports* 3: 195-202.

Jordan, A., and P. Reichard. 1998. Ribonucleotide reductases. *Annual review of biochemistry* 67: 71-98.

Kaper, James B., James P. Nataro, and Harry LT Mobley. 2004. Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* 2: 123-140.

Karmali, M. A., M. Petric, C. Lim, P. C. Fleming, and B. T. Steele. 1983. Sporadic cases of hemolytic uremic syndrome associated with fecal cytotoxin and cytotoxin producing *Escherichia coli* in stools. *The Lancet i* 619: 620.

Karmali, Mohamed A. 2004. Prospects for Preventing Serious Systemic Toxemic Complications of Shiga Toxin Producing *Escherichia coli* Infections Using Shiga Toxin Receptor Analogues. *Journal of Infectious Diseases* 189: 355-359.

Katoh, Kazutaka, Kazuharu Misawa, Kei-ichi Kuma, and Takashi Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research* 30: 3059-3066.

Kearse, Matthew, Richard Moir, Amy Wilson, Steven Stones-Havas, Matthew Cheung, Shane Sturrock, Simon Buxton et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647-1649.

Keen, James E., and Robert O. Elder. 2002. Isolation of shiga-toxigenic *Escherichia coli* O157 from hide surfaces and the oral cavity of finished beef feedlot cattle. *Journal of the American Veterinary Medical Association* 220: 756-763.

Keir, Lindsay S., Stephen D. Marks, and Jon Jin Kim. 2012. Shiga toxin-associated hemolytic uremic syndrome: current molecular mechanisms and future therapies. *Drug design, development and therapy* 6: 195.

Kim, Kwang-Pyo, Yannick Born, Rudi Lurz, Fritz Eichenseher, Markus Zimmer, Martin J. Loessner, and Jochen Klumpp. 2012. Inducible *Clostridium perfringens* bacteriophages phiS9 and phiS63: Different genome structures and a fully functional sigK intervening element. *Bacteriophage* 2: 89-97.

Klumpp, Jochen, Rob Lavigne, Martin J. Loessner, and Hans-Wolfgang Ackermann. 2010. The SPO1-related bacteriophages. *Archives of virology* 155: 1547-1561.

Klumpp, Jochen, Richard Calendar, and Martin J. Loessner. 2010b. Complete nucleotide sequence and molecular characterization of *Bacillus* phage TP21 and its relatedness to other phages with the same name. *Viruses* 2: 961-971.

Kong, Minsuk, Minsik Kim, and Sangryeol Ryu. 2012. Complete genome sequence of *Bacillus cereus* bacteriophage PBC1. *Journal of virology* 86: 6379-6380.

Kropinski, Andrew M., Melissa Hayward, M. Dorothy Agnew, and Ken F. Jarrell. 2005. The genome of BCJA1c: a bacteriophage active against the alkaliphilic bacterium, *Bacillus clarkii*. *Extremophiles* 9: 99-109.

Labont, Jessica M., Karen E. Reid, and Curtis A. Suttle. 2009. Phylogenetic analysis indicates evolutionary diversity and environmental segregation of marine podovirus DNA polymerase gene sequences. *Applied and environmental microbiology* 75: 3634-3640.

Lee, Ju-Hoon, Hakdong Shin, Bokyung Son, and Sangryeol Ryu. 2012. Complete genome sequence of *Bacillus cereus* bacteriophage BCP78. *Journal of virology* 86: 637-638.

Lee, Ju-Hoon, Hakdong Shin, Bokyung Son, Sunggi Heu, and Sangryeol Ryu. 2013. Characterization and complete genome sequence of a virulent bacteriophage B4 infecting food-borne pathogenic *Bacillus cereus*. *Archives of virology* 158: 2101-2108.

Lee, Young-Duck, and Jong-Hyun Park. 2010. Genomic sequence of temperate phage 250 isolated from emetic *B. cereus* and cloning of putative endolysin. *Food Science and Biotechnology* 19: 1643-1648.

Li, Linlin, Tongling Shan, Chunlin Wang, Colette Ct, John Kolman, David Onions, Frances MD Gulland, and Eric Delwart. 2011. The fecal viral flora of California sea lions. *Journal of virology* 85: 9909-9917.

Lindsay, Jodi A., Alexey Ruzin, Hope F. Ross, Natasha Kurepina, and Richard P. Novick. The gene for toxic shock toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. *Molecular microbiology* 29: 527-543.

Loeffler, Jutta M., Daniel Nelson, and Vincent A. Fischetti. 2001. Rapid killing of *Streptococcus pneumoniae* with a bacteriophage cell wall hydrolase. *Science* 294: 2170-2172.

Loessner, Martin J., Simon K. Maier, Helmut Daubek-Puza, Gunther Wendlinger, and Siegfried Scherer. 1997. Three *Bacillus cereus* bacteriophage endolysins are unrelated but reveal high homology to cell wall hydrolases from different bacilli. *Journal of bacteriology* 179: 2845-2851.

Lorenz, Laura, Bridget Lins, Jonathan Barrett, Andrew Montgomery, Stephanie Trapani, Anne Schindler, Gail E. Christie, Steven G. Cresawn, and Louise Temple. 2013. Genomic characterization of six novel *Bacillus pumilus* bacteriophages. *Virology* 444: 374-383.

Loretz, Marianne, Roger Stephan, and Claudio Zweifel. 2011. Antibacterial activity of decontamination treatments for cattle hides and beef carcasses. *Food Control* 22: 347-359.

Lozupone, Catherine, and Rob Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and environmental microbiology* 71: 8228-8235.

Marchler-Bauer, Aron, Shennan Lu, John B. Anderson, Farideh Chitsaz, Myra K. Derbyshire, Carol DeWeese-Scott, Jessica H. Fong et al. 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research* 39: D225-D229.

Mayer, Melinda J., John Payne, Michael J. Gasson, and Arjan Narbad. 2010. Genomic sequence and characterization of the virulent bacteriophage phiCTP1 from *Clostridium tyrobutyricum* and heterologous expression of its endolysin. *Applied and environmental microbiology* 76: 5415-5422.

McCann, Joshua C., Tryon A. Wickersham, and Juan J. Loor. 2014. High-throughput methods redefine the rumen microbiome and its relationship with nutrition and metabolism. *Bioinformatics and biology insights* 8: 109.

Meessen-Pinard, Mathieu, Ognjen Sekulovic, and Louis-Charles Fortier. 2012. Evidence of in vivo prophage induction during *Clostridium difficile* infection. *Applied and environmental microbiology* 78: 7662-7670.

Meijer, Wilfried JJ, Jos A. Horcajadas, and Margarita Salas. 2001. phi29 family of phages. *Microbiology and Molecular Biology Reviews* 65: 261-287.

Minakhin, Leonid, Ekaterina Semenova, Jing Liu, Anatoly Vasilov, Elena Severinova, Tarasii Gabisonia, Ross Inman, Arcady Mushegian, and Konstantin Severinov. 2005. Genome sequence and gene expression of *Bacillus anthracis* bacteriophage Fah. *Journal of molecular biology* 354: 1-15.

Morales, Cesar A., Brian B. Oakley, Johnna K. Garrish, Gregory R. Siragusa, Mary B. Ard, and Bruce S. Seal. 2012. Complete genome sequence of the podoviral bacteriophage phiCP24R, which is virulent for *Clostridium perfringens*. *Archives of virology* 157: 769-772.

Morohoshi, Tomohiro, Yoshiaki Tominaga, Nobutaka Someya, and Tsukasa Ikeda. 2012. Complete genome sequence and characterization of the N-acylhomoserine lactone-degrading gene of the potato leaf-associated *Solibacillus silvestris*. *Journal of bioscience and bioengineering* 113: 20-25.

Moumen, Bouziane, and Alexei Sorokin. 2012. Sequence analysis of inducible prophage phIS3501 integrated into the haemolysin II gene of *Bacillus thuringiensis var israelensis* ATCC35646. *Genetics research international* 543286.

Nale, Janet Y., Jinyu Shan, Peter T. Hickenbotham, Warren N. Fawley, Mark H. Wilcox, and M. R. Clokie. 2012. Diverse temperate bacteriophage carriage in *Clostridium difficile* 027 strains. *PLoS One* 7: e37263.

Nariya, Hirofumi, Shigeru Miyata, Eiji Tamai, Hiroshi Sekiya, Jun Maki, and Akinobu Okabe. 2011. Identification and characterization of a putative endolysin encoded by episomal phage phiSM101 of *Clostridium perfringens*. *Applied microbiology and biotechnology* 90: 1973-1979.

Nataro, James P., and James B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clinical microbiology reviews* 11: 142-201.

Nordlund, Pr., and Reichard, Peter. 2006. Ribonucleotide reductases. *Annu. Rev. Biochem* 75: 681-706.

O'Brien, Alison D., John W. Newland, Steven F. Miller, Randall K. Holmes, H. Williams Smith, and Samuel B. Formal. 1984. Shiga-like toxin-converting phages from *Escherichia coli* strains that cause hemorrhagic colitis or infantile diarrhea. *Science* 226: 694-696.

O'Flaherty, Sarah, R. Paul Ross, and Aidan Coffey. 2009. Bacteriophage and their lysins for elimination of infectious bacteria. *FEMS microbiology reviews* 33: 801-819.

Oakley, Brian B., Eldin Talundzic, Cesar A. Morales, Kelli L. Hiett, Gregory R. Siragusa, Nikolay V. Volozhantsev, and Bruce S. Seal. 2011. Comparative genomics of four closely related *Clostridium perfringens* bacteriophages reveals variable evolution among core genes with therapeutic potential. *BMC genomics* 12: 282.

Ohya, Tatsuo, Masato AKIBA, and Hiroya ITO. 2001. Use of a Trial Probiotic Product in Calves Experimentally Infected with *Escherichia coli* O157. *Japan Agric Res Quart* 35: 189-194.

Overbeek, Ross, Tadhg Begley, Ralph M. Butler, Jomuna V. Choudhuri, Han-Yu Chuang, Matthew Cohoon, Valrie de Crcy-Lagard et al. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic acids research* 33: 5691-5702.

Paddock, Zachary, Xiaorong Shi, Jianfa Bai, and T. G. Nagaraja. 2012. Applicability of a multiplex PCR to detect O26, O45, O103, O111, O121, O145, and O157 serogroups of *Escherichia coli* in cattle feces 1. *Veterinary microbiology* 156: 381-388.

Paton, James C., and Paton, Adrienne W. 1998. Pathogenesis and diagnosis of Shiga toxin-producing *Escherichia coli* infections. *Clinical microbiology reviews* 11: 450-479.

Peterson, R. E., Terry J. Klopfenstein, Galen E. Erickson, J. Folmer, S. Hinkley, Rodney A. Moxley, and David R. Smith. 2007. Effect of *Lactobacillus acidophilus* strain NP51 on *Escherichia coli* O157: H7 fecal shedding and finishing performance in beef feedlot cattle. *Journal of Food Protection* 70: 287-291.

Poss, Bjrn, Lieven De Zutter, Marc Heyndrickx, and Lieve Herman. 2008. Novel differential and confirmation plating media for Shiga toxin-producing *Escherichia coli* serotypes O26, O103, O111, O145 and sorbitol-positive and-negative O157. *FEMS microbiology letters* 282: 124-131.

Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. 2005. Epidemiology of *Escherichia coli* O157: H7 outbreaks, United States, 1982-2002. *Emerging Infectious Diseases* 11: 603-609

Redondo, Rodrigo AF, Anne Kupczok, Gertraud Stift, and Jonathan P. Bollback. 2013. Complete genome sequence of the novel phage MG-B1 Infecting *Bacillus weihenstephanensis*. *Genome announcements* 1: e00216-13.

Rice, William C., Michael L. Galyean, Stephen B. Cox, Scot E. Dowd, and N. Andy Cole. 2012. Influence of wet distillers grains diets on beef cattle fecal bacterial community structure. *BMC microbiology* 12: 25.

Richardson, Susan E., Mohamed A. Karmali, Laurence E. Becker, and Charles R. Smith. 1988. The histopathology of the hemolytic uremic syndrome associated with verocytotoxin-producing *Escherichia coli* infections. *Human pathology* 19: 1102-1108.

Rigobelo, E. E. C., N. Karapetkov, S. A. Maest, F. A. vila, and D. McIntosh. 2014. Use of probiotics to reduce faecal shedding of Shiga toxin-producing *Escherichia coli* in sheep. *Beneficial microbes* 6: 53-60.

Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, Hebert RJ, Olcott ES, Johnson LM, Hargrett NT. 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *New England Journal of Medicine* 308: 681-685.

Ross, Elizabeth M., Steve Petrovski, Peter J. Moate, and Ben J. Hayes. 2013. Metagenomics of rumen bacteriophage from thirteen lactating dairy cattle. *BMC microbiology* 13: 242.

Russell, James B., and Jennifer L. Rychlik. 2001. Factors that alter rumen microbial ecology. *Science* 292: 1119-1122.

Sakowski, Eric G., Erik V. Munsell, Mara Hyatt, William Kress, Shannon J. Williamson, Daniel J. Nasko, Shawn W. Polson, and K. Eric Wommack. 2014. Ribonucleotide reductases reveal novel viral diversity and predict biological and ecological features of unknown marine viruses. *Proceedings of the National Academy of Sciences* 111: 15786-15791.

Sandvig, Kirsten, and Van Deurs, Bo. 1996. Endocytosis, intracellular transport, and cytotoxic action of Shiga toxin and ricin. *Physiological Reviews* 76: 949-966.

Sakaguchi, Yoshihiko, Tetsuya Hayashi, Ken Kurokawa, Keisuke Nakayama, Kenshiro Oshima, Yukako Fujinaga, Makoto Ohnishi, Eiichi Ohtsubo, Masahira Hattori, and Keiji Oguma. 2005. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. *Proceedings of the National Academy of Sciences of the United States of America* 102: 17472-17477.

Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, Jones JL, Griffin PM. 2011. Foodborne illness acquired in the United States' major pathogens. *Emerg Infect Dis* 17.

Schloss, Patrick D., Sarah L. Westcott, Thomas Ryabin, Justine R. Hall, Martin Hartmann, Emily B. Hollister, Ryan A. Lesniewski et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75: 7537-7541.

Schmidt, Helen F., Eric G. Sakowski, Shannon J. Williamson, Shawn W. Polson, and K. Eric Wommack. 2014. Shotgun metagenomics indicates novel family A DNA polymerases predominate within marine virioplankton. *The ISME journal* 8: 103-114.

Schmitz, Jonathan E., Anu Daniel, Mattias Collin, Raymond Schuch, and Vincent A. Fischetti. 2008. Rapid DNA library construction for functional genomic and metagenomic screening. *Applied and environmental microbiology* 74: 1649-1652.

Schoenfeld, Thomas, Melodee Patterson, Paul M. Richardson, K. Eric Wommack, Mark Young, and David Mead. 2008. Assembly of viral metagenomes from yellowstone hot springs. *Applied and environmental microbiology* 74: 4164-4174.

Scholl, Dean, and Carl Merril. 2005. The genome of bacteriophage K1F, a T7-like phage that has acquired the ability to replicate on K1 strains of *Escherichia coli*. *Journal of bacteriology* 187: 8499-8503.

Schuch, Raymond, and Vincent A. Fischetti. The secret life of the anthrax agent *Bacillus anthracis*: bacteriophage-mediated ecological adaptations. *PLoS One* 4: e6532.

Schuch, R., A. J. Pelzek, S. Kan, and V. A. Fischetti. 2010. Prevalence of *Bacillus anthracis*-like organisms and bacteriophages in the intestinal tract of the earthworm *Eisenia fetida*. *Applied and environmental microbiology* 76: 2286-2294.

Seal, Bruce S., Derrick E. Fouts, Mustafa Simmons, Johnna K. Garrish, Robin L. Kuntz, Rebekah Woolsey, Kathleen M. Schegg, Andrew M. Kropinski, Hans-W. Ackermann, and Gregory R. Siragusa. 2011. *Clostridium perfringens* bacteriophages phiCP39O and phiCP26F: genomic organization and proteomic analysis of the virions. *Archives of virology* 156: 25-35.

Shanks, Orin C., Catherine A. Kelty, Shawn Archibeque, Michael Jenkins, Ryan J. Newton, Sandra L. McLellan, Susan M. Huse, and Mitchell L. Sogin. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. *Applied and environmental microbiology* 77: 2992-3001.

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13: 2498-2504.

Sheehan, Michelle M., Jos L. Garca, Rubens Lpez, and Pedro Garca.1997. The lytic enzyme of the *pneumococcal* phage Dp-1: a chimeric lysin of intergeneric origin. *Molecular microbiology* 25: 717-725.

Shin, Hakdong, Ju-Hoon Lee, Jaeeun Park, Sunggi Heu, and Sangryeol Ryu. 2014. Characterization and genome analysis of the *Bacillus cereus*-infecting bacteriophages BPS10C and BPS13. *Archives of virology* 159: 2171-2175.

Short, Cindy M., and Curtis A. Suttle. 2005. Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Applied and environmental microbiology* 71: 480-486.

Smeesters, Pierre R., Pierre-Alexandre Drze, Sabrina Bousbata, Kaarle J. Parikka, Sophie Timmery, Xiaomin Hu, David Perez-Morga et al. 2011. Characterization of a novel temperate phage originating from a cereulide-producing *Bacillus cereus* strain. *Research in microbiology* 162: 446-459.

Sozhamannan, Shanmuga, Michael McKinstry, Shannon M. Lentz, Matti Jalasvuori, Farrell McAfee, Angela Smith, Jason Dabbs et al. 2008. Molecular characterization of a variant of *Bacillus anthracis*-specific phage AP50 with improved bacteriolytic activity. *Applied and environmental microbiology* 74: 6792-6796.

Sperandio, Vanessa. 2010. SdiA sensing of acyl-homoserine lactones by enterohemorrhagic *E. coli* (EHEC) serotype O157: H7 in the bovine rumen. *Gut microbes* 1: 432-435.

Stromberg, Zachary R., Nicholas W. Baumann, Gentry L. Lewis, Nicholas J. Sevart, Natalia Cernicchiaro, David G. Renter, David B. Marx, Randall K. Phebus, and Rodney A. Moxley. 2015. Prevalence of Enterohemorrhagic *Escherichia coli* O26, O45, O103, O111, O121, O145, and O157 on Hides and Preintervention Carcass Surfaces of Feedlot Cattle at Harvest. *Foodborne Pathogens and Disease* 12: 631-638.

Suttle, Curtis A. 2007. Marine viruses major players in the global ecosystem. *Nature Reviews Microbiology* 5: 801-812.

Suzek, Baris E., Hongzhan Huang, Peter McGarvey, Raja Mazumder, and Cathy H. Wu. 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23: 1282-1288.

Suzuki, Motoshi, Shonen Yoshida, Elinor T. Adman, A. Blank, and Lawrence A. Loeb. 2000. *Thermus aquaticus* DNA Polymerase I Mutants with Altered Fidelity Interactions Mutations in the O-helix. *Journal of Biological Chemistry* 275: 32728-32735.

Stanford, K., T. P. Stephens, and T. A. McAllister. 2011. Use of model super-shedders to define the role of pen floor and hide contamination in the transmission of O157: H7. *Journal of animal science* 89: 237-244.

Stavric, S., B. Buchanan, and T. M. Gleeson. 1992. Competitive exclusion of *Escherichia coli* O157: H7 from chicks with anaerobic cultures of faecal microflora. *Letters in applied microbiology* 14: 191-193.

Tabor, Stanley, and Charles C. Richardson. 1987. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proceedings of the National Academy of Sciences* 84: 4767-4771.

Tabor, Stanley, and Charles C. Richardson. 1995. A single residue in DNA polymerases of the *Escherichia coli* DNA polymerase I family is critical for distinguishing between deoxy-and dideoxyribonucleotides. *Proceedings of the National Academy of Sciences* 92: 6339-6343.

Tarr, Phillip I., Carrie A. Gordon, and Wayne L. Chandler. 2005: Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *The Lancet* 365: 1073-1086.

Thingstad, T. Frede. 2000. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography* 45: 1320-1328.

Thomas, Julie A., Stephen C. Hardies, Mandy Rolando, Shirley J. Hayes, Karen Lieman, Christopher A. Carroll, Susan T. Weintraub, and Philip Serwer. 2007. Complete genomic sequence and mass spectrometric analysis of highly diverse, atypical *Bacillus thuringiensis* phage 0305836. *Virology* 368: 405-421.

Uchiyama, Jumpei, Mohammad Rashel, Yoshihiro Maeda, Iyo Takemura, Shigeyoshi Sugihara, Kazue Akechi, Asako Muraoka, Hiroshi Wakiguchi, and Shigenobu Matsuzaki. 2008. Isolation and characterization of a novel *Enterococcus faecalis* bacteriophage phiEF24C as a therapeutic candidate. *FEMS Microbiology letters* 278: 200-206.

Vge, Selina, Julia E. Storesund, and T. Frede Thingstad. 2013. Adding a cost of resistance description extends the ability of virus-host model to explain observed patterns in structure and function of pelagic microbial communities. *Environmental microbiology* 15: 1842-1852.

Van Donkersgoed, Joyce, Tom Graham, and Vic Gannon. 1999. The prevalence of verotoxins, *Escherichia coli* O157: H7, and *Salmonella* in the feces and rumen of cattle at processing. *The Canadian Veterinary Journal* 40: 332.

van Elsas, Jan Dirk, Alexander V. Semenov, Rodrigo Costa, and Jack T. Trevors. 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *The ISME journal* 5: 173-183.

van Elsas, Jan Dirk, Mario Chiurazzi, Cyrus A. Mallon, Dana Elhottova, Vclav Kritufek, and Joana Falco Salles. 2012. Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proceedings of the National Academy of Sciences* 109: 1159-1164.

Van Houdt, Rob, Abram Aertsen, Pieter Moons, Kristof Vanoirbeek, and Chris W. Michiels. 2006. N-acyl-L-homoserine lactone signal interception by *Escherichia coli*. *FEMS microbiology letters* 256: 83-89.

Vzquez-Baeza, Yoshiki, Meg Pirrung, Antonio Gonzalez, and Rob Knight. 2013. EMPeror: a tool for visualizing high-throughput microbial community data. *Structure* 585: 20.

Vzquez-Boland, Jos A., Gustavo Domnguez-Bernal, Bruno Gonzlez-Zorn, Jrgen Kreft, and Werner Goebel. 2001. Pathogenicity islands and virulence evolution in *Listeria*. *Microbes and Infection* 3: 571-584.

Vivant, Anne-Laure, Dominique Garmyn, Pierre-Alain Maron, Virginie Nowak, and Pascal Piveteau. 2013. Microbial diversity and structure are drivers of the biological barrier effect against *Listeria monocytogenes* in soil. *PloS one* 8: e76991.

Vold, L., A. Holck, Y. Wasteson, and H. Nissen. 2000. High levels of background flora inhibits growth of *Escherichia coli* O157: H7 in ground beef. *International Journal of Food Microbiology* 56: 219-225.

Volozhantsev, Nikolay V., Vladimir V. Verevkin, Vasily A. Bannov, Valentina M. Krasilnikova, Vera P. Myakinina, Eugeni L. Zhilenkov, Edward A. Svetoch, Norman J. Stern, Brian B. Oakley, and Bruce S. Seal. 2011. The genome sequence and proteome of bacteriophage phiCPV1 virulent for *Clostridium perfringens*. *Virus research* 155: 433-439.

Volozhantsev, Nikolay V., Brian B. Oakley, Cesar A. Morales, Vladimir V. Verevkin, Vasily A. Bannov, Valentina M. Krasilnikova, Anastasia V. Popova et al. 2012. Molecular characterization of podoviral bacteriophages virulent for *Clostridium perfringens* and their comparison with members of the Picovirinae. *PloS one* 7: e38283.

Wagner, Patrick L., Jonathan Livny, Melody N. Neely, David WK Acheson, David I. Friedman, and Matthew K. Waldor. 2002. Bacteriophage control of Shiga toxin 1 production and release by *Escherichia coli*. *Molecular microbiology* 44: 957-970.

Winget, Danielle M., and K. Eric Wommack. 2009. Diel and daily fluctuations in virioplankton production in coastal ecosystems. *Environmental microbiology* 11: 2904-2914.

Winter, Christian, Thierry Bouvier, Markus G. Weinbauer, and T. Frede Thingstad. 2010. Trade-offs between competition and defense specialists among unicellular

planktonic organisms: the "killing the winner" hypothesis revisited. *Microbiology and Molecular Biology Reviews* 74: 42-57.

Wommack, K. Eric, Jaysheel Bhavsar, Shawn W. Polson, Jing Chen, Michael Dumas, Sharath Srinivasiah, Megan Furman, Sanchita Jamindar, and Daniel J. Nasko. 2012. VIROME: a standard operating procedure for analysis of viral metagenome sequences. *Standards in genomic sciences* 6: 427.

Wommack, K. Eric, Daniel J. Nasko, Jessica Chopyk, and Eric G. Sakowski. Counts and sequences, observations that continue to change our understanding of viruses in nature. *Journal of Microbiology* 53: 181-192.

Yasbin, Ronald E., and Frank E. Young. 1974. Transduction in *Bacillus subtilis* by bacteriophage SPP1. *Journal of Virology* 14: 1343-1348.

Zhao, Tong, Michael P. Doyle, Barry G. Harmon, Cathy A. Brown, PO Eric Mueller, and Andrew H. Parks. 1998. Reduction of carriage of enterohemorrhagic *Escherichia coli* O157: H7 in cattle by inoculation with probiotic bacteria. *Journal of clinical microbiology* 36: 641-647.

Zhao, Tong, Suzana Tkalcic, Michael P. Doyle, Barry G. Harmon, Cathy A. Brown, and Ping Zhao. 2003. Pathogenicity of enterohemorrhagic *Escherichia coli* in neonatal calves and evaluation of fecal shedding by treatment with probiotic *Escherichia coli*. *Journal of Food Protection* 66: 924-930.

Zhao, L., P. J. Tyler, J. Starnes, C. L. Bratcher, D. Rankins, T. A. McCaskey, and Luxin Wang. 2013. Correlation analysis of Shiga toxin-producing *Escherichia coli* shedding and faecal bacterial composition in beef cattle. *Journal of applied microbiology* 115: 591-603.

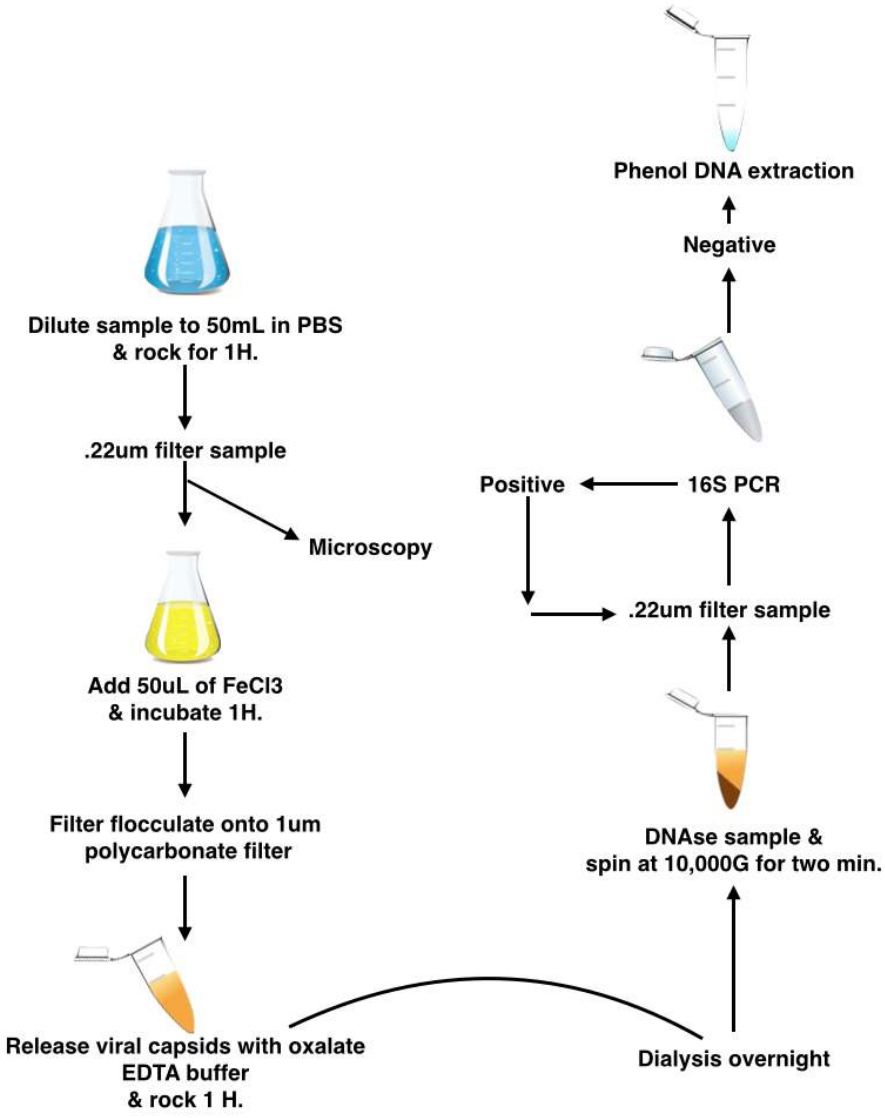
Zhu, W., Lomsadze, A., and Borodovsky, M. 2010. Ab initio gene identification in metagenomic sequences. *Nucleic acids research* 38: e132-e132.

Zimmer, Markus, Siegfried Scherer, and Martin J. Loessner. 2002. Genomic analysis of *Clostridium perfringens* bacteriophage phi3626, which integrates into guaA and possibly affects sporulation. *Journal of bacteriology* 184: 4359-4368.

Appendix A

PHAGE ISOLATION PROTOCOL

An abbreviated protocol for viral particle isolation.



Appendix B

PHENOL CHLORIDE CRACK METHOD FOR DNA ISOLATIONL

1. Add 1 volume of Phenol Chloroform and 1/10 volume of 3M NaOAc to sample.

2. Spin at 14,000 rpm, room temperature for 10 minutes.

3. Transfer supernatant to a clean eppendorf tube (this contains your DNA)

i. You may back extract to try and get more DNA by adding 200 µl of Elution Buffer to the bottom layer and spinning again at 14,000 rpm, room temperature, for 10 minutes.

ii. The supernatant can then be transferred to the eppendorf tube from the first spin. 4. Add an equal volume of chloroform to the supernatant and spin at 14,000 rpm, room temperature, for 10 minutes to remove residual phenol.

5. Transfer supernatant to a new clean eppendorf tube, this is you clean DNA.

6. Add 2 volumes of 100% EtOH to the supernatant and incubated at least 24 hours at -20°.

7. Spin at 14,000 rpm, 4°for 30 minutes to an hour (DNA pellet will be visible at the bottom of the tube).

8. Decant ethanol, careful not to touch the pellet.

9. Add 300 µl of fresh 70% EtOH and spin at 14,000 rpm, room temperature for 10 minutes.

10. Decant off all EtOH and allow pellet to dry for roughly 5 minutes (careful not to over dry the pellet).

11. Re-suspend in 20-50 µl of Elution Buffer.