UNIVERSITY OF CALIFORNIA,
IRVINE

Evolution of bacteriophages infecting *Enterococcus* from the human microbiome

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Molecular Biology and Biochemistry

by

Stephen Joseph Wandro

Dissertation Committee:
Assistant Professor Katrine Whiteson, Chair
Professor Jennifer Martiny
Associate Professor Ilhem Messaoudi

2019

DEDICATION

To my parents, Bob and Cathy

Thank you for your love and support

And to my partner, Linh Anh

Let's write the next chapter together

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

None of this work would have been possible without the support and guidance of many people. To my parents: thank you for being a constant source of love and support. You set me up for success and set examples for the kind of person I want to be. Thank you to my advisor, Katrine Whiteson, who has been an amazing mentor and an endless source of fun ideas. Thank you for teaching me how to be a scientist and allowing me the freedom to pursue my interests. I will miss our lab dinners, Halloween murals, and Julefrokost celebrations. Thank you to my committee member Jennifer Martiny for helping me dive into the world of phages and experimental evolution. To my other committee member, Ilhem Messaoudi, thank you for your thoughtful advice and feedback.

To all the past and present members of the Whiteson Lab, thank you for making these last five years so much fun. You have been a great group of people to work with and this work wouldn't be possible without the support of the whole lab. Andrew, Whitney, Tara, and Joann, you are my valued colleagues. Thank you for all the help you have given me over the years and for all the good times at Whiteson Lab happy hours. Thanks to my undergraduate researchers, Bushra and Cyril, for being excellent helpers. To Clark, you are the hardest working individual I have ever met and your passion for science inspires me. Thank you for being my lab hands. Thank you Claudia for your advice and help in the lab. Thank you to the people of microbial group for providing me good discussions and good happy hours.

Thank you to my wonderful partner Linh Anh. You are the best thing to come out of my graduate school experience. I am continually impressed by you as a scientist, a writer, a science communicator, and as a person. You give me support when I need it and are always there with me to celebrate all the big and little things.

Financial support was provided by the Center for Virus Research at UC Irvine. Thank you to the Mortazavi lab for generously allowing us to use their sequencer.

Chapter 1 was published with the permission of American Society for Microbiology. The text of this chapter is a reprint of the material as it appears in the mSphere. Chapter 2 was published with the permission of American Society for Microbiology. The text of this chapter is a reprint of the material as it appears in the journal mSystems. Chapter 3 was published with the permission of the Frontiers Media. The text of this chapter is a reprint of the material as it appears in the journal Frontiers in Microbiology.

CURRICULUM VITAE

Stephen J. Wandro
Contact: SteveWandro@gmail.com
GitHub: swandro

Education

PhD ..... September 2014 – June 2019
University of California, Irvine
Department: Molecular Biology and Biochemistry
Thesis adviser: Katrine Whiteson
GPA: 4.0

Master's Degree ..... September 2014 – November 2017
University of California, Irvine
Department: Molecular Biology and Biochemistry
GPA: 4.0

Bachelor of Science ..... June 2014
University of California, Los Angeles
Major: Microbiology, Immunology, and Molecular Genetics
GPA: 3.86

Skills

- Proficient in programming in Python, R, Bash
- Next generation DNA sequencing and analysis
- Advanced molecular techniques
- Scientific communication
- Multivariate statistics
- Genomics, metagenomics, metabolomics, transcriptomics analyses

Lab Experience

Whiteson Lab (UCI) ..... Spring 2015 – June 2019
Earned a PhD investigating bacteria and phage communities in the human gut. My research involved using next-generation sequencing approaches and experimental evolution to study how bacteria and phage from the human gut co-evolve.

Kohn Lab (UCLA) ..... Fall 2011 – Summer 2014
As an undergraduate researcher, I contributed to a project for using lentiviral vectors for gene therapy to treat ADA linked Severe Combined Immunodeficiency.

Training

- Woods Hole Microbial Diversity Course ..... Summer 2017
- West Coast Metabolomics Center workshop at UC Davis ..... Summer 2015
- Research Mentoring Training at UC Irvine ..... Spring 2016
- UC Irvine's National Short Course in Systems Biology ..... January 2016
- Phage Hunters Project course series at UCLA ..... Winter-Spring 2014

Awards and Fellowships

- NIH T32 training grant "Molecular Biology of Eukaryotic Viruses" (2 years of funding)
- Edward K. Wagner Memorial Award in Virology 2017
- Magna Cum Laude at UCLA
- Dean's Honors List for 9 quarters at UCLA
- National Merit Scholarship winner in 2010

Publications

Wandro S, Carmody L, Gallagher T, LiPuma JJ, Whiteson K. 2017. *mSystems*.
Making It Last: Storage Time and Temperature Have Differential Impacts on Metabolite Profiles of Airway Samples from Cystic Fibrosis Patients.
DOI: 10.1128/mSystems.00100-17.

Wandro S, Osborne S, Enriquez C, Bixby C, Arrieta A, Whiteson K. 2018 *mSphere*.
The Microbiome and Metabolome of Preterm Infant Stool Are Personalized and Not Driven by Health Outcomes, Including Necrotizing Enterocolitis and Late-Onset Sepsis.
DOI: 10.1128/mSphere.00104-18.

Wandro S, Oliver A, Gallagher T, Weihe C, England W, Martiny JBH, Whiteson K. 2018. *Frontiers in Microbiology, Virology*. Predictable molecular adaptation of coevolving Enterococcus faecium and lytic phage EfV12-phi1.
DOI: 10.3389/fmicb.2018.03192

Research Presentations

- Research talk at North American Microbiome Congress, Washington D.C. 2019
- Poster presented at Lake Arrowhead Microbial Genomics (won poster award) 2018
- Research talk at American Society for Microbiology Summer 2017.
- Research talk at California Microbiome Meeting 2017
- Poster presented at UCI ICTS Research Day 2015, 2016, and 2017.
- Poster presented at Multi-omics for Microbiomes Conference September 2015
- Poster presented at Children's Hospital Orange County Research Week 2015

Work/Volunteer History

Mentoring undergraduate researchers.....Fall 2015 – Present
Help undergraduate students at UC Irvine with independent research projects.

Ran journal club for first-year students .....Spring 2016
Helped a group of first-year PhD students prepare for their preliminary exam.

Stanford Hoover Institute student archivist .....Summer 2011 – 2012
Assisted researchers and performed clerical work in archives

ABSTRACT OF THE DISSERTATION

Evolution of bacteriophages infecting *Enterococcus* from the human microbiome

By

Stephen Wandro

Doctor of Philosophy in Molecular Biology & Biochemistry

University of California, Irvine, 2019

Assistant Professor Katrine Whiteson, Chair

Enterococcus can be both a friend and a foe in the human microbiome. As an opportunistic pathogen, *Enterococcus* is normally benign. However, *Enterococcus* is responsible for many hospital-acquired infections and has shown rising rates of vancomycin resistance. Bacteriophages (phages) could be an alternative to antibiotics to target these antibiotic resistant bacteria, but the evolutionary and phenotypic outcomes of phage-bacteria interactions need to be investigated more thoroughly.

We begin by investigating the gut microbiome of preterm infants that had been exposed to antibiotics to learn about the scenarios in which *Enterococcus* blooms occur in the gut (Chapter 1). We show that antibiotic exposed gut microbiomes are dominated by facultative anaerobes such as *Enterococcus* and *Enterobacteriaceae*. Further, we show that the bacterial composition is correlated to the overall metabolite profile. Metabolomics is a powerful tool for investigating microbial metabolism, and we contribute to the effort of developing standardized practices for metabolomics in the human microbiome by showing that freezing microbial samples required for long term storage (Chapter 2).

Next, we investigated how phages could be used as therapeutics for treating *Enterococcus* blooms and infections (Chapters 3, 4, and 5). When *Enterococcus* is grown with its phages *in vitro*, it evolves resistance to phage infection by mutating exopolysaccharide synthesis genes. These mutations alter the exopolysaccharides on the surface of the bacterial cell to prevent binding of phage. Further, this mechanism of resistance appears to be a general mechanism leading to resistance against a diverse array of phages. This work demonstrates that experimental evolution is a powerful tool for characterizing interactions between bacteria and phages.

Phage therapy is often administered as a cocktail of multiple phages, but there are no rules or best practices described for combining phages to be most effective. We show that *in vitro*, *Enterococcus* phage cocktails are more effective at preventing the growth of phage-resistant mutants, but the composition of the cocktail is important. Genetically diverse phage cocktails performed better than cocktails of related phages. My work demonstrates some of the outcomes of phage-*Enterococcus* interactions and will move us closer to applying phage therapy to treat *Enterococcus* infections.

INTRODUCTION

The human microbiome

The human microbiome contains incredible microbial diversity (Yatsunenko et al. 2012; Human Microbiome Project Consortium 2012). From birth, we are colonized by a dense community of microbes that profoundly impact our health and development (Rosa et al. 2014). Recently, affordable next-generation sequencing technologies have allowed us to probe these microbial communities on a large scale for the first time. We can now see that the human microbiome is comprised of a complex community including bacteria, viruses, and fungi (Lloyd-Price et al. 2017). Interactions among microbes and host cells form a complex web that we are only beginning to untangle.

Most of the microbes associated with the human body live in the gastrointestinal tract, specifically in the large intestines. The animal gut is one of the densest communities of microbes on the planet, containing 10^{11} bacteria per gram of feces (Sender, Fuchs, and Milo 2016). The ratio of human to bacterial cells in the human body is nearly 1:1 by current estimates, but, the genetic diversity encoded by microbes far exceeds the genetic diversity within the human genome (Human Microbiome Project Consortium 2012). The human genome encodes for approximately 20,000 genes, while the bacteria in our microbiome encode millions of genes (Qin et al. 2010). Therefore, much of what makes us unique individuals is encoded by our microbes.

Microbes in the gastrointestinal tract perform many beneficial functions for the host, including aiding in digestion, preventing colonization of pathogens, producing nutrients, and training the immune system (Pflughoeft and Versalovic 2012). Microbiome interactions with the immune system have become a center of focus as microbes are

implicated in many health conditions including obesity, diabetes, Crohn's Disease, and autism (Turnbaugh et al. 2006; Larsen et al. 2010; Buffington et al. 2016; Hsiao et al. 2013). However, one of the clearest messages that is emerging from human microbiome research is that each person's microbiome is unique. Since, there is no universal "healthy" microbiome, it has been a challenge to link microbiome states to health and disease.

Metabolomics to study the microbiome

In addition to sequencing, metabolomics has emerged as a powerful tool for probing microbial systems. Metabolomics is the study of small molecules called metabolites. It can be used to study the metabolism of microbes and provide insight into the functional role of the microbial community that cannot be gained from simply studying bacterial abundances or genomes (Lamichhane et al. 2018b; Wandro, Osborne, et al. 2017). Metabolomics has been used in microbiome studies to find biomarkers of health and disease (Sévin et al. 2015). Many of the ways our body interacts with the microbes in the gut is through the metabolites they produce. For example, production of short-chain fatty acids by gut microbes through the fermentation of fiber has been shown to benefit gut barrier function, modulate the immune system, and improve glucose homeostasis (Chang et al. 2014; Morrison and Preston 2016). There are many ways to perform metabolomics and each has its benefits and drawbacks (Lamichhane et al. 2018a). Practices and techniques for metabolomics in microbiome studies are much less standardized than genomics, making it difficult to compare results across studies (Wandro, Carmody, et al. 2017). In addition, the gut metabolome is highly personalized just like the microbiome (Wandro, Osborne, et al. 2017).

Phages in the microbiome

Bacteriophages are the viruses that infect bacteria. They are abundant in the microbiome and contribute greatly to the diversity and personalized nature of the microbiome. The human gut virome consists of almost entirely bacteriophages (phages) that are lytic predators or lysogens of the resident bacterial community (Ogilvie and Jones 2015; Minot et al. 2011). Little is known about how the phages affect the abundance and composition of bacteria in humans (Manrique et al. 2016). Like bacterial communities, phage communities in the healthy adult gut are mostly stable over time, but some classes of phages demonstrate such high mutation rates that speciation has been observed within an individual over only 2.5 years (Minot et al. 2011, 2013). Phages have also been shown to facilitate horizontal gene transfer between bacteria, which could result in spreading antibiotic resistance genes (Canchaya et al. 2003; Keen et al. 2017). Despite these important functions of phages, few studies have examined the bacterial and phage components of the microbiome together over time (Reyes et al. 2015; Minot et al. 2011; Lim et al. 2015; Breitbart et al. 2008).

Microbiome in early development

The personalized nature of the microbiome makes it difficult but important to understand the forces that influence its initial assembly and development after birth. Early life is thought to be a critical time of microbiome development that has lifelong health consequences (Gibson et al. 2016; Koenig et al. 2011; C. Palmer et al. 2007; Kostic et al. 2015). Initial exposure to microbes occurs during birth as the infant microbiome is seeded by microbes from the mother and from the environment (Asnicar et al. 2017; Ferretti et al. 2018; Brooks et al. 2017). Many factors have been shown to alter the assembly of the infant

microbiome, including diet, prematurity, and antibiotic exposure (Pannaraj et al. 2017; Schulfer and Blaser 2015; Bokulich et al. 2016; Gibson et al. 2016). Antibiotic exposure in particular can severely disrupt the microbiome and result in the permanent loss of bacterial strains (Dethlefsen and Relman 2011a).

It is generally thought that it takes about three years for a mature microbiome to develop (Yatsunenko et al. 2012). In the first few years of life, critical interactions occur between microbes in the gut and immune cells that help to train the immune system to correctly recognize threats (Cox et al. 2014; Thaiss et al. 2016). Therefore, having the correct composition of microbes in the gut during those first years is essential for the development of the immune system. The gut microbiome of infants in developed countries is increasingly diverging from infants in other parts of the world due to changes including diet, sanitation, family size, and antibiotic usage (Yatsunenko et al. 2012; Charbonneau et al. 2016; Henrick et al. 2018). Infants, and especially premature infants, are prescribed antibiotics at a higher rate than any other age group (Ting et al. 2016). The microbiomes of antibiotic-exposed infants are often dominated by fast-growing, facultative anaerobes, such as *Enterococcus*, *Streptococcus*, and *Enterobacteriaceae* (Sommer and Dantas 2011; Dethlefsen and Relman 2011a; Ubeda et al. 2010a; Bäumler and Sperandio 2016). Overgrowth of normally low-abundance members of the infant gut could have lifelong immune consequences (Zeng, Inohara, and Nuñez 2017; Fujimura et al. 2016).

Antibiotics perturb the microbiome

Antibiotics are the main tool used to treat bacterial infections, but are relatively non-specific and will kill many commensal microbes in the process (Dethlefsen et al. 2008). After the human microbiome is decimated by antibiotics, a different composition may

reemerge upon regrowth of the community (Dethlefsen and Relman 2011a). After antibiotic administration, the gut is temporarily more aerobic, allowing the quick growth of facultative anaerobes (Sjlund et al. 2003; Dethlefsen and Relman 2011a). This can lead to a loss of diversity over time or even severe illness in the case of *Clostridium difficile* infections (Buffie et al. 2012). A more targeted approach to treating bacterial infections and dysbioses would be beneficial to minimize adverse effects on the microbiome.

Overuse of antibiotics in the last several decades has also led to rising rates of antibiotic resistant bacterial infections (Bradford 2018). Using antibiotics to kill bacteria selects for bacteria that are resistant; the more an antibiotic is used, the more common resistance to that antibiotic will become. Thus, in every instance when a new antibiotic is introduced, antibiotic resistance is quickly observed (Zaman et al. 2017). This becomes a major problem when bacteria evolve resistance to multiple antibiotics and we are no longer able to treat those infections (Nikaido 2009). Additionally, antibiotic resistant bacteria tend to spread around hospitals, infecting people with weakened immune systems and diminished capability to fight off infections (D. L. Smith et al. 2004).

Enterococcus

Enterococcus is a genus of gram-positive bacteria that are present at low abundance in the intestines of most humans. The two most common species of *Enterococcus* found associated with humans is *Enterococcus faecalis* and *Enterococcus faecium* (Agudelo Higuita and Huycke 2014a). *Enterococcus* has an ancient association with animals, and has been present in the microbiomes of humans and our ancestors for over 400 million years (Francois Lebreton, Willems, and Gilmore 2014). However, *Enterococcus* is an opportunistic pathogen in humans, accounting for 9% of hospital acquired infections

(Magill et al. 2014). They are commonly responsible for UTIs, wound infections, root canal infections, endocarditis, and sepsis (Koch et al. 2004). The *Enterococcus* genus is naturally resistant to many antibiotics, and vancomycin resistant *Enterococcus* has emerged as a major health crisis (Agudelo Higuita and Huycke 2014b). Alternative therapies are needed to deal with the rising rates of vancomycin-resistant *Enterococcus*.

Phage therapy

Using phages to treat bacterial infections is known as phage therapy (Kortright et al. 2019b). Phages target a much narrower range of bacteria than antibiotics, resulting in less harmful effects to the microbiota (Kortright et al. 2019b). Phage therapy has been utilized since the 1920s, but was largely abandoned after the discovery of antibiotics (Summers 2012). With the supply of effective antibiotics dwindling, phage therapy could provide an alternative or complementary option for treating multidrug resistant bacterial infections. Phage therapy has been successful in a few recent cases, and interest in developing phage therapy to treat antibiotic resistant infections has been increasing (Chan et al. 2018; Duplessis et al. 2017). Phages have even shown potential in treating *Enterococcus in vitro* and in mice (Khalifa et al. 2016, 2015a; Chatterjee et al. 2019). However, more work must be done to isolate diverse phages and characterize their interactions with *Enterococcus* before phage therapy can be a viable option.

Phages and bacteria constantly co-evolve, but the extent and implications of this coevolution in the human microbiome is not known. The dynamics and outcomes of phage-bacteria coevolution have been studied extensively in model systems, including *Escherichia*, *Pseudomonas*, and *Synechococcus* (Hall et al. 2011; Paterson et al. 2010; Marston et al. 2012; Perry et al. 2015). In laboratory experimental evolution, arms-race

dynamics are normally observed in which reciprocal evolution of bacterial resistance and phage infectivity lead to more generally resistant bacterial phenotypes and more generally infectious phage phenotypes (Mizoguchi et al. 2003a; Martiny et al. 2014; Marston et al. 2012). The evolutionary and phenotypic outcomes of co-evolution between bacteria and phages will be important to understand before phage therapy can be widely adopted.

Bacteria can evolve resistance to phage infection by several routes including restriction modification, abortive infection, CRISPR-Cas, and blocking of adsorption (Dy et al. 2014). Blocking adsorption is commonly seen in laboratory evolution experiments, and occurs by modification, deletion, or blocking of the binding target/receptor (Meyer et al. 2012; Mizoguchi et al. 2003b). Mutations allowing phage resistance often have an associated fitness cost such as impaired growth, colonization defects, or increased antibiotic susceptibility (Scanlan, Buckling, and Hall 2015; B. Koskella et al. 2012; Chan et al. 2016; Chatterjee et al. 2019; Lennon et al. 2007). These fitness costs would allow phage therapy to be effective even if bacteria evolve resistance to phage. A method of preventing the evolution of bacterial resistance to phage is to use cocktails of multiple phages (Nale et al. 2018a; M. Yen, Cairns, and Camilli 2017a; Kortright et al. 2019a). There are currently no guidelines or best practices for combining phages for phage cocktails. Diverse phages that utilize different receptors may reduce the likelihood that a single mutation will provide broad resistance, but this has not been thoroughly tested across multiple systems (R. C. T. Wright et al. 2018; Flores et al. 2011a; Lennon et al. 2007).

Goals and scope of dissertation

The aim of this dissertation is to elucidate the evolutionary interactions between human associated *Enterococcus* and its bacteriophages to move the field closer to utilizing

phage therapy to treat *Enterococcus*. First, we investigate *Enterococcus* blooms in the microbiomes of preterm infants that have been exposed to antibiotics using both genomic and metabolomic approaches. Next, we begin to investigate the evolutionary interactions with phage by showing the *in vitro* co-evolution of *Enterococcus* with one of its lytic bacteriophages. We then expand that view to an entire sub-family of *Enterococcus* bacteriophages. Finally, we investigate the principles of how to combine bacteriophages for optimal killing of *Enterococcus*. We hope this work will lay the groundwork for understanding the outcomes of co-evolution between *Enterococcus* and its phages.

CHAPTER 1

The microbiome and metabolome of pre-term infant stool is personalized, and not driven by health outcomes including necrotizing enterocolitis and late-onset sepsis

Co-authors: Stephanie Osborne, Claudia Enriquez, Christine Bixby, Antonio Arrieta, and Katrine Whiteson

ABSTRACT

The assembly and development of the gut microbiome in infants has important consequences for immediate and long-term health. Preterm infants represent an abnormal case for bacterial colonization because of early exposure to bacteria and frequent use of antibiotics. To better understand the assembly of the gut microbiota in preterm infants, fecal samples were collected from 32 very low birthweight preterm infants over the first six weeks of life. Infant health outcomes included healthy, late-onset sepsis, and necrotizing enterocolitis (NEC). We characterized the bacterial composition by 16S rRNA gene sequencing and metabolome by untargeted gas chromatography mass spectrometry. Preterm infant fecal samples lacked beneficial *Bifidobacterium* and were dominated by *Enterobacteriaceae*, *Enterococcus*, and *Staphylococcus* due to the near uniform antibiotic administration. Most of the variance between the microbial community compositions could be attributed to which baby the sample came from (Permanova R^2=0.48, $p<0.001$), while clinical status (healthy, NEC, or late-onset sepsis), and overlapping time in the NICU did not explain a significant amount of variation in bacterial composition. Fecal metabolomes were also found to be unique to the individual (Permanova R^2=0.43, $p<0.001$) and weakly associated with bacterial composition (Mantel statistic $r = 0.23 \pm 0.05$, $p<0.05$). No

measured metabolites were found to be associated with necrotizing enterocolitis, late-onset sepsis or a healthy outcome. Overall, preterm infants gut microbial communities were personalized and reflected antibiotic usage.

IMPORTANCE

Preterm infants face health problems likely related to microbial exposures including sepsis and necrotizing enterocolitis. However, the role of the gut microbiome in preterm infant health is poorly understood. Microbial colonization differs from healthy term babies because it occurs in the NICU and is often perturbed by antibiotics. We measured bacterial compositions and metabolomic profiles of 77 fecal samples from thirty-two preterm infants to investigate the differences between microbiomes in health and disease. Rather than finding microbial signatures of disease, we found the preterm infant microbiome and metabolome were both personalized, and that the preterm infant gut microbiome is enriched in microbes that commonly dominate in the presence of antibiotics. These results contribute to the growing knowledge of the preterm infant microbiome and emphasize that a personalized view will be important to disentangle the health consequences of the preterm infant microbiome.

INTRODUCTION

Early life exposure to microbes and their metabolic products is a normal part of development, with enormous and under-explored impact on the immune system. The intestinal microbiota of infants initially assembles by exposure to the mother's microbiota and microbes in the environment (Gibson et al. 2016; Rosa et al. 2014; Bäckhed et al. 2015; Bokulich et al. 2016). In healthy breast-fed infants, *Bifidobacteria longum* spp. infantis capable of digesting human-milk oligosaccharides dominate the infant gut (Frese et al.

2017). When infants are born preterm, they are exposed to environmental and human associated microbes earlier in their development than normal, and rarely harbor *Bifidobacteria* spp. in their gut communities. We do not yet understand the effects of altering the timing of initial bacterial exposure. With numerous emerging health consequences related to the microbiome, understanding factors that influence this initial assembly of the microbiome will be important.

Preterm infants are routinely treated with antibiotics, enriching for microbes that can colonize in the presence of antibiotics (Bokulich et al. 2016; Nobel et al. 2015; Cox et al. 2014). While antibiotics have tremendously reduced infant mortality, their effect on microbiota assembly and resulting health consequences is not fully understood. Prenatal and postnatal antibiotics have been shown to reduce the diversity of the infant intestinal microbiota (Tanaka et al. 2009; Keski-Nisula et al. 2013). Children under two years old are prescribed antibiotics at a higher rate than any other age group, and 85% of extremely low birthweight infants (< 1000 g) are given at least one course of antibiotics (Ting et al. 2016). Even if an infant is not exposed to antibiotics after birth, approximately 37% of pregnant women use antibiotics over the course of the pregnancy (Stokholm et al. 2013).

Perturbing the microbiota of infants can have immediate and long-lasting health consequences. In the short term, infants can be infected by pathogenic bacteria that results in sepsis, which is categorized as early-onset or late-onset depending on the timing after birth. Preterm infants are also at high risk to develop necrotizing enterocolitis (NEC), which is a devastating disease that causes portions of the bowel to undergo necrosis. NEC is one of the leading causes of mortality in preterm infants, who make up 90% of NEC cases (Fitzgibbons et al. 2009). The incidence of NEC among low birthweight preterm infants is

approximately 7% and causes death in about one third of cases. The exact causes of NEC are not known, but an excessive inflammatory response to intestinal bacteria may be involved (Nanthakumar et al. 2011).

Many of the long-term consequences of microbial colonization are believed to be mediated by interactions between the intestinal microbiota and the immune system. In addition to direct interactions, the microbiota interacts with the immune system through the production of metabolites that can be taken up directly by immune and epithelial cells (Dodd et al. 2017a; Wikoff et al. 2009). For example, bacterial production of short chain fatty acids can affect health and integrity of the intestinal epithelia and immune cells (Willemsen et al. 2003; Chang et al. 2014; C. J. Kelly et al. 2015). However, few studies use metabolites alongside bacterial community profiling. In fact, the healthy composition of an infant fecal metabolome is understudied.

In this retrospective study, we follow the changes in the gut microbiota over time in 32 very low birth weight (< 1500 grams) preterm infants born at Children's Hospital Orange County. We simultaneously track the bacterial composition and metabolite profile over time. Infants were classified into three groups based on health outcomes: healthy, late-onset sepsis, and NEC. The composition of the intestinal microbiota was measured by 16S rRNA gene sequencing of fecal samples taken over time. Preterm infant guts were dominated by *Enterobacteriaceae* and *Enterococcus*, and *Staphylococcus*. Untargeted metabolomics analysis of the fecal samples by gas chromatography mass spectrometry (GC-MS) revealed a personalized metabolome that was weakly associated with the bacterial composition.

RESULTS

Patient cohort

A total of 77 fecal samples were collected from 32 very low birth weight infants in the NICU at Children's Hospital Orange County from 2011 to 2014 (**Table 1.1, Figure 1.1**). Birthweights ranged from 620 – 1570 grams. Fecal samples were collected between day 7 and 75 of life. Sampling time and number of fecal samples varied. Three or more longitudinal samples were available from ten of the infants, while one or two samples were available from the remaining 22 infants. Three infants developed NEC, eight developed late-onset sepsis, and 21 remained healthy. Twelve infants were delivered vaginally while the remaining 22 were delivered by cesarean section. All infants were fed by either breastmilk or a combination of breastmilk and formula. Twenty-four infants had record of receiving antibiotics at some point during the sampling period, the most common being ampicillin and gentamycin.

Microbial Community Characterization

We sequenced the 16S rRNA gene content of each fecal sample to determine bacterial composition. The total bacterial load of each fecal sample was measured by qPCR of the 16S rRNA gene and scaled to the total weight of stool that DNA was extracted from. Among all infants, bacterial abundances vary over four orders of magnitude and were lower in infants that developed NEC or late-onset sepsis ($p < 0.001$) (F**igure 1.2**). The high variation in bacterial load is likely due to the near uniform use of antibiotics. Bacterial communities were composed of mostly *Enterobacteriaceae*, *Enterococcus*, *Staphylococcus*, and *Bacteroides* (**Figure 1.2**). Most samples were dominated by one to three genera of bacteria. Only three infants (two fed breastmilk, one fed breastmilk and formula) were colonized at

greater than 1% relative abundance by *Bifidobacteria*, which emerging evidence suggests is a key member of the infant microbiome. However, we note that the primers used are able to detect 1741 out of 5146 (30%) of *Bifidobacteria* species represented in the Ribosome Project Database including 38 *infantis* substrains, versus 2177663 out of 3196041 (68%) of all bacterial species in the database (Cole et al. 2014). No single bacterial OTU or community composition was consistently found for infants that became sick (NEC or late-onset sepsis) compared to the infants that remained healthy.

Longitudinal sampling revealed that over the course of days, the bacterial composition could change dramatically (**Figure 1.3**). Permutational Multivariate Analysis of Variance (PERMANOVA) was applied to determine which of the known clinical factors explained the most variance in the bacterial community composition. The individual explained 48% ($p < 0.001$) of the variance in the samples, meaning that about half of the total variance among all tested fecal samples could be attributed to the infant the fecal sample came from (**Supplemental table 1.1a**). Delivery mode explained a smaller proportion of variance (12% variance, $p < .05$), but none of the other factors explained a significant amount of variation in the bacterial composition, including infant health, overlapping dates in the NICU, or feeding mode. Only vaginally born infants were colonized by *Bacteroides* (four out of nine infants) while none of the twenty-two infants born by C-section were colonized. Four of the infants in the study are twins. Twin set 1 (infants 12 and 13) had a similar microbial composition while the other three sets did not (**Supplemental Figure 1.1**). Diversity of the bacterial communities was low, as expected for preterm infants. Alpha diversity as measured by Shannon index increased overall with age, but the trend was not significant (linear model $R^2 = 0.005$, $p = 0.52$) (**Figure 1.4a**). Other clinical factors including

health outcome, feeding (breastmilk versus breastmilk and formula), antibiotic use, and delivery mode were tested for an effect on the alpha diversity (**Figure 1.4b-e**). None of the factors were associated with a difference in alpha diversity except recorded antibiotic use, in which Shannon diversity was unexpectedly lower on average in infants that did not have a record of receiving antibiotics (Wilcoxon rank sum test p=0.06). It should be noted that although six infants did not have a record of antibiotic use, records may be incomplete due to hospital transfers.

Metabolomics

Metabolite profiles of infant fecal samples were analyzed by gas chromatography mass spectrometry, which measures small primary metabolites. Over 400 small molecules were detected from each fecal sample and 224 metabolites were known compounds. Metabolites were grouped into the following categories: amino acid metabolism, bile acids, central metabolism, fatty acids, fermentation products, lipid metabolism, nucleotide metabolism, organic acids, sterols, sugars, sugar acids, sugar alcohols, and vitamin metabolism (**Figure 1.5, Supplemental table 1.2**). No metabolites or categories of metabolites were found to be associated with necrotizing enterocolitis or late-onset sepsis. The metabolite profile of each infant was seen to vary over time, similar to the amount of variation seen in the bacterial composition (**Figure 1.6**). PERMANOVA analysis to determine which factors explain the most variance in the metabolite profile indicate that the individual explains 43% ($p < 0.001$) of the variation (**Supplemental table 1.1b**).

To determine which metabolites might be useful for tracking bacterial metabolism in the infant gut, we examined metabolites with consistent abundance among infants versus those that varied (**Supplemental figure 1.2**). In general, sugars, central metabolites, and

amino acids were variable while fatty acids, sterols, organic acids, and bile acids were more consistent. Infant 23, which developed necrotizing enterocolitis at day 16 of life, had low abundances of amino acid metabolites the two days prior to disease onset (**Figure 1.5**). However, several of the healthy control infants also had similarly low abundances of amino acid metabolites. The individual signal of each infant's metabolome is far more evident than any trends due to clinical factors (**Supplemental table 1.1b**).

Bacterial composition associated with metabolite profile

Bacterial metabolism in the gut is expected to contribute to the abundances of metabolites detected in fecal samples. We wanted to know if fecal samples with a similar bacterial composition were also similar in their metabolite profile. We employed a Mantel test using Pearson correlations between distances among bacterial compositions of samples and distances among metabolite profiles of samples. Because bacterial compositions and metabolite profiles are personalized, using multiple samples from a single infant would skew the result. Therefore, one sample from each infant was randomly selected 100 times to remove the effect of the individual and the Mantel test was applied to each subset. The average Mantel statistic of $r = 0.23 \pm 0.05$ ($p < 0.05$) indicates a weak but significant association between the bacterial composition and metabolite profile. Also, within individual infants, shifts in the bacterial composition are accompanied by shifts in the metabolome. Infants 17, 23, and 31 have dramatic shifts in both bacterial composition and metabolome profile over time, while infants 10 and 37 remain stable in both bacterial composition and metabolome.

To investigate the correlations driving this overall association, we calculated correlations between bacterial abundances and metabolite intensities (**Figure 1.7a**). *Staphylococcus*

had the most positive correlations including several classes of sugar metabolites, organic acids, and central metabolites. Fatty acids, lipid metabolism, and amino acids were positively correlated with the commonly abundant gut colonizers *Enterobacteriaceae* and *Bacteroides*, and negatively correlated with the commonly low abundance colonizers *Staphylococcus* and *Enterococcus*. We also looked more specifically at individual metabolites correlated with bacterial abundances (**Figure 1.7b**). Bacteroidetes were found to be positively correlated with succinate ($r = 0.85$). Many other weak correlations ($r < 0.5$) exist between bacterial abundances and metabolite intensities, but the sample size is not large enough to distinguish signal from noise.

DISCUSSION

Bacterial compositions in this cohort were consistent with the emerging picture from other studies that show the preterm infant gut harbors communities dominated by facultative anaerobes including *Enterobacteriaceae*, *Enterococcus*, and *Staphylococcus* (Gibson et al. 2016; Rosa et al. 2014; Grier et al. 2017). These communities appear to be enriched in commonly antibiotic resistant organisms (Sommer and Dantas 2011). While we expected to find associations between bacterial community composition and health outcome in this cohort, we were surprised to find that there were not clear signatures of microbiome composition linked to NEC or sepsis. In larger cohorts, associations between particular bacteria or metabolites with NEC have been reported, however, they are not universal signatures across patients, and may reflect patient variation more than disease etiology (Morrow et al. 2013; Sim et al. 2015; Heida et al. n.d.; Cassir et al. n.d.). In fact, the strongest signal in both the microbiome and metabolome data from this cohort was the infant from whom the sample was taken. Overall, preterm infant microbiomes in this study were

shaped by antibiotics, which have a strong impact on all patients, regardless of health outcome.

Although the bacterial composition of infant guts varied over time, we saw longitudinal samples from individual infants remained highly personalized over several weeks; nearly half of the variation in the microbial community compositions can be explained by which individual the sample came from. The stability of animal-associated microbiomes is an active area of research, with studies finding that the individual microbiome of an adult remains stable through time (Faith et al. 2013), but can be perturbed by extreme changes in diet or antibiotics (Dethlefsen and Relman 2011a; David et al. 2013; Brüssow 2016). The bacterial composition in the adult gut largely returns to its previous state one month after antibiotic treatment, but altering the initial assembly of the microbiota in infants can have long lasting health consequences (Dethlefsen and Relman 2011a; Dethlefsen et al. 2008; Cox et al. 2014; Schulfer and Blaser 2015). Previous work has found ampicillin and gentamycin (the most common antibiotics taken by infants in this study) to have an inconsistent effect on bacterial diversity, sometimes increasing and sometimes decreasing diversity (Gibson et al. 2016). Similarly, in these infants, ampicillin and gentamycin resulted in more variation in bacterial diversity, but there was no clear trend of increasing or decreasing diversity. However, antibiotics change the dominant members of the microbiota which could have profound effects on immune development and growth (Metsälä et al. 2015; Mueller et al. 2015; Cox et al. 2014; Schulfer and Blaser 2015). Evidence is emerging that a healthy infant gut microbiota is dominated by *Bifidobacteria* with the capacity to digest human milk oligosaccharides in breastmilk (Karav et al. 2016; Frese et al. 2017; Underwood et al. 2017). The lack of a core *Bifidobacteria* community in

infants could leave the microbiota open to colonization by facultative anaerobes like we observed in these infants (Stewart et al. 2016). Proteobacteria such as *Enterobacteriaceae* are commonly seen to increase in abundance after antibiotic administration (Sommer and Dantas 2011; Andersen et al. 2016; Ubeda et al. 2010a). Indeed, infants in this study had microbiomes that were shaped by antibiotic use. Although six of the thirty-two infants in this study did not have recorded antibiotic use around sampling time, the microbiota can still be affected by prenatal antibiotics taken by the mother (Clark et al. 2006; Cox et al. 2014; Schulfer and Blaser 2015).

Microbiome studies have become widespread, so that a typical bacterial composition is well characterized in a range of sample cohorts. However, the same cannot be said for the metabolome. There is a dearth of knowledge about what a consensus healthy infant fecal metabolome should be, making comparisons for the cohort in this study difficult. To add to the complexity, each metabolomic approach detects subsets of metabolites, and depends on sample extraction and other method choices. Increasing the frequency of metabolomics data collection in microbiome studies would be a huge step forward for the field. Baseline knowledge about the typical connections between metabolite abundances and bacterial metabolism should be collected, so that molecules that have consistent abundances in a healthy state could give context to data generated from clinical samples in different disease states.

Untargeted metabolomics can survey many metabolites in a biological sample to provide a snapshot of the active metabolism in a system such as the human gut. Metabolite profiles of preterm infants in this study were found to be personalized to a similar degree as the bacterial composition. This is in contrast to a previous study on full term infants that

showed the metabolomic profile to be stable, and weakly associated with bacterial composition, over the first few years of life (Kostic et al. 2015). Personalized metabolic signatures of disease hold great promise to complement microbiota profiling in human systems (Stewart et al. 2016; C. J. Kelly et al. 2015). However, analyzing metabolomic data is challenging because an array of inputs contribute to the abundances of metabolites in fecal samples including bacterial metabolism, host biology, and food intake.

We report a number of correlations between bacteria and metabolites in preterm infant feces, and bacterial metabolism has been previously shown to contribute to metabolite abundances in humans and mice (Wikoff et al. 2009; S. Yen et al. 2015; Dodd et al. 2017b). Short chain fatty acids are now commonly measured and associated with bacterial fermentation in the gut (Morrison and Preston 2016). In this study, the only short-chain fatty acid detected was succinate, which we found to be correlated with the presence of *Bacteroides*, which produces acetate and succinate from carbohydrate fermentation (Miller 1978). We also detected several medium-chain fatty acids, which were generally correlated with the abundance of *Bacteroides* and *Enterobacteriaceae*. None of the twenty-two C-section born infants in this study were colonized by *Bacteroides*, possibly due to a lack of vertical transmission from the mother during birth (Bäckhed et al. 2015).

Overall, we find preterm infant microbiomes are shaped by shared exposures especially to antibiotics, leading to the dominance of antibiotic resistant facultative anaerobes such as *Enterococcus* spp.. The anaerobic, milk degrading *Bifidobacteria* were largely absent, even in preterm infants with access to breastmilk, possibly due to a lack of exposure to microbes from family members in the sterile hospital environment along with antibiotics. Our

understanding of the health consequences of microbial colonization under these antibiotic-enriched circumstances is still in its infancy.

MATERIALS AND METHODS

Sample Collection

Stool samples from diapers of preterm infants in the neonatal intensive care unit at Children's Hospital Orange County were collected by nurses over three years from 2011 to 2014. Samples were immediately stored at -20 °C then transferred to -80 °C no more than three days post-collection. Samples were kept at -80 °C and thawed once for DNA extraction and metabolomics. A total of 77 stool samples were collected from 32 preterm infants.

DNA extraction and 16S rRNA gene sequencing

Stool samples were thawed once and DNA was extracted from 10 mg using a Zymo Fecal DNA MiniPrep Kit (#D6010). The V3 and V4 region of the 16SrRNA gene was amplified with two-stage PCR. The first PCR amplified the V3 to V4 region of the 16S rRNA gene using 341F and 805R primers: forward primer (5'- CCTACGGGNGGCWGCAG-3') and reverse primer (5'- GACTACHVGGGTATCTAATCC -3') (Herlemann et al. 2011). These primers also added an overhang so that barcodes and Illumina adaptors could be added in the second PCR. The first PCR was done as follows: 30 cycles of 95 °C 30 seconds; 65 °C 40 seconds; 72 °C 1 minute. Immediately after completion of the first PCR, primers with sample specific barcodes and Illumina adapter sequences were added and a second PCR was performed as follows: 9 cycles 94 °C for 30 seconds; 55 °C 40 seconds; 72 °C 1 minute. PCR reactions were cleaned using Agencourt AMPure XP magnetic beads (#A63880) using the recommended protocol. Amplicons were run on an agarose gel to confirm amplification

and then pooled. The amplicon pool was run on an agarose gel and the 500bp fragment was cut out and gel extracted using Millipore Gel Extraction Kit (#LSKGEL050). The sequencing library was quantified using Quant-iT Pico Green dsDNA Reagent and sent to Laragen Inc. for sequencing on the Illumina MiSeq platform with 250bp paired-end reads producing a total of 2.4 million paired-end reads.

qPCR for bacterial load

The bacterial load of each fecal sample was measured with quantitative PCR of a conserved region of the 16S gene. The following primers were used: (5'- TCC TAC GGG AGG CAG CAG T-3'), (5'- GGA CTA CCA GGG TAT CTA ATC CTG TT-3'). PerfeCTa SYBER Green SuperMix Reaction Mix (Quantabio #95054) was used to quantify DNA from samples. Relative abundance of 16S rRNA genes relative to the mass of stool was compared for each sample. Total fecal DNA was measured with Quant-iT Pico Green dsDNA Assay Kit (ThermoFisher #P11496).

Sequence processing

Sequences were quality filtered using PrinSeq to remove adapters as well as any sequences less than 120 base-pairs, containing any ambiguous bases, or with a mean PHRED quality score of less than 30 (Schmieder and Edwards 2011). Reads were found to drop steeply in quality after 140 base pairs, so all reads were trimmed to 140 base pairs. The forward read contained the V3 region in the high quality first 140 base pairs, while the V4 region was captured in the low-quality region of the reverse reads. Therefore, we used only the forward reads for subsequent analyses.

Bacterial community analysis

Quantitative Insights Into microbial Ecology (QIIME) was used for de novo OTUs picking using the Swarm algorithm with a clustering threshold of 8 (Caporaso et al. 2010; Mahé et al. 2014). This resulted in 2,810 OTUs among all samples. OTUs containing only one sequence were filtered out, leaving 212 OTUs. Taxonomy was assigned to each OTU using QIIME and the Greengenes 13_8 database. An OTU table was constructed and used for downstream analysis. Ten rarefactions were performed on the OTU table down to 2000 reads per sample, which was the largest number of reads that allowed retention of most samples. QIIME was used to calculate alpha diversity by Shannon index and beta diversity by the average weighted UniFrac distance of the ten rarefactions. Community composition barplots, Principal Coordinate Analysis (PCoA) plots, and alpha diversity plots were created using R and the ggplot2 package (Wickham 2009; Lozupone and Knight 2005). All R scripts are included in the supplemental information.

Untargeted metabolomics by GC-MS

When fecal samples were thawed for DNA extraction, approximately 50 mg was collected and refrozen at -80 ° for metabolomics. Samples were sent on dry ice to the West Coast Metabolomics Center (WCMC) at UC Davis for untargeted metabolomics by gas chromatography time-of-flight mass spectrometry. Metabolites were extracted from fecal samples with a 3:3:2 mixture of isopropanol, acetonitrile, and water respectively before derivatization and GC-MS analysis by Fiehn Lab standard operating procedures (Kind et al. 2009, 2017; Cajka and Fiehn 2016). Metabolite profiles were compared by calculating Manhattan distances between samples based on all metabolite intensities and visualized by PCoA using the vegan and ape packages in R (Paradis, Claude, and Strimmer 2004; Oksanen et al. 2016).

Permutational multivariate analysis of variance (PERMANOVA)

PERMANOVA was used to determine factors that explained variance in bacterial community composition and metabolite profile. PERMANOVA was performed using the adonis function in the vegan package in R. The input for PERMANOVA was a UniFrac distance matrix of the 16S data and Manhattan distances of the metabolite profiles. Briefly, PERMANOVA quantifies the variation among samples explained by the given groupings compared to randomized groupings. To measure the variance explained by individual infant, we excluded samples that had fewer than three longitudinal samples, leaving ten infants. To measure the variance explained by health outcome, we again only included infants with three or more longitudinal samples and groups were permuted among infants, not samples, so that the effect of the individual will be minimal. When performing PERMANOVA for factors other than individual, we accounted for the longitudinal sampling by averaging samples from each individual.

Correlations between bacteria and metabolites

Pearson correlations between bacterial abundances and normalized metabolite intensities were calculated using the cor function in R. Correlations were calculated between the relative abundances of all bacterial classes and all metabolite intensities among all samples in all infants. Only the four most highly abundant general of bacteria were used to ensure no results were skewed by taxa present in only one or a few samples. For each class of metabolite, the average of all correlations between metabolites in that class and each taxon was calculated so that trends between bacterial taxa and classes of metabolites could be visualized by heatmap.

Mantel test

To determine if fecal samples with similar bacterial compositions also have similar metabolite profiles, a Mantel test was performed. To account for the effect of longitudinal sampling, each dataset was randomly subsampled down to one sample per infant. A Bray-Curtis dissimilarity matrix was computed for the bacterial composition and Manhattan distances calculated for metabolite intensities. The Mantel function in the vegan package of R was used to calculate the Mantel statistic for a Pearson correlation between the two dissimilarity matrices. The average and standard deviation of the Mantel statistic r and p-value for the 100 Mantel tests was reported.

DATA AVAILABILITY

Raw sequence data is available on the SRA at accession: SRP137076. OTU tables, raw metabolomics data, a markdown file of sequence processing workflow, and R scripts used for analyses are available at https://github.com/swandro/preterm_infant_analysis.

ACKNOWLEDGEMENTS

This project was supported by a UCI Single Investigator grant, the CORCL SIIG-2014-2015-51, a pilot grant from the UC Davis West Coast Metabolomics Core as part of NIH-DK097154, and start-up funds for the Whiteson Lab in UC Irvine's Molecular Biology and Biochemistry Department. Thanks to CORCL for funding.

Thank you to Ying (Lucy) Lu for help with qPCR experiments. Thank you Celine Mouginot and Adam Martiny for providing 16S primers and protocols. Thanks to Ilhem Messaoudi and members of the Whiteson lab for editing and discussion. Thanks to Megan Showalter and Prof. Oliver Fiehn of the UC Davis West Coast Metabolomics Core were very helpful in carrying out untargeted GC-MS profiling.

FIGURES AND TABLES

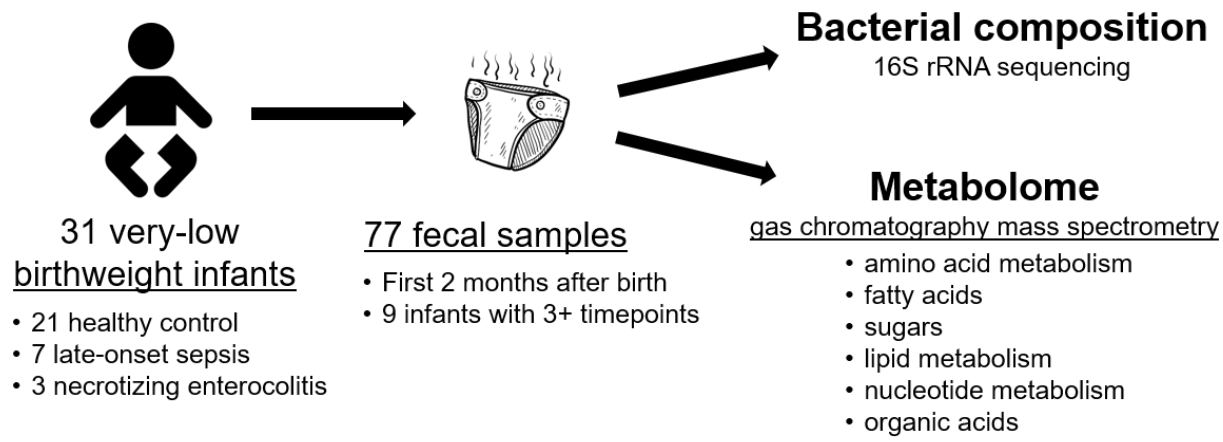


Figure 1.1 Study design schematic. Longitudinal fecal samples were collected over the first 75 days of life from very low birthweight infants in the NICU. Bacterial compositions and metabolomes were characterized.

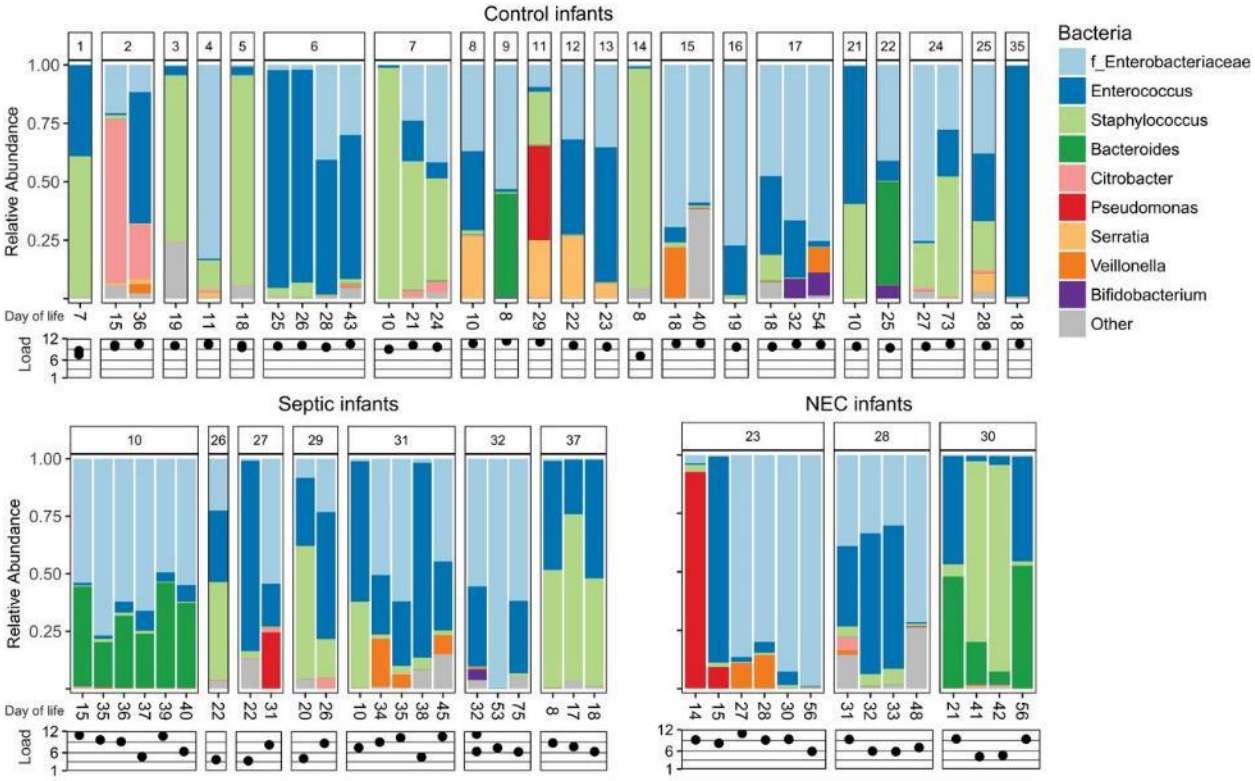


Figure 1.2 Bacterial composition and bacterial load of preterm infant guts. Stacked barplots show relative abundance of bacteria at the genus level in all infant samples. The family Enterobacteriaceae is included because genus level resolution was not available. Log scaled relative bacterial load is shown underneath each sample.

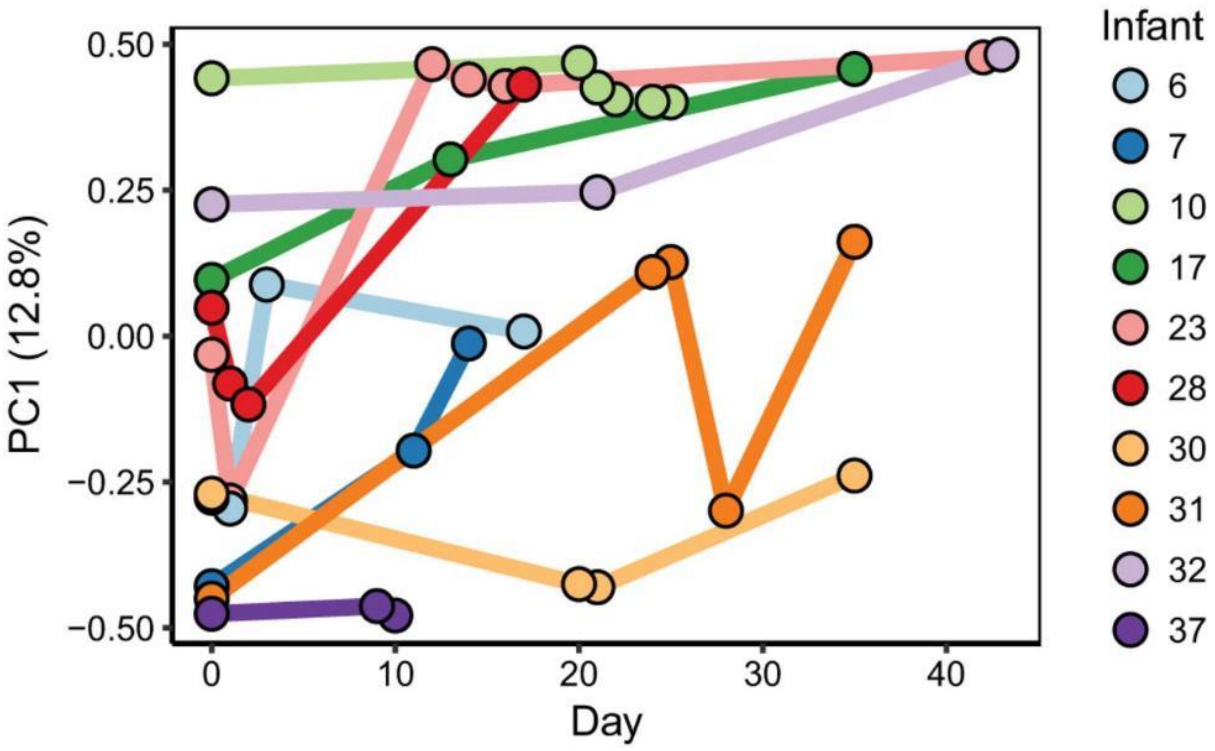


Figure 1.3 First axis of PCoA based on weighted uniFrac distances between bacterial communities plotted over time. Each dot represents a single fecal sample and is colored by infant. Lines connect samples for each infant to show change over time. Only infants with three or more longitudinal samples shown.

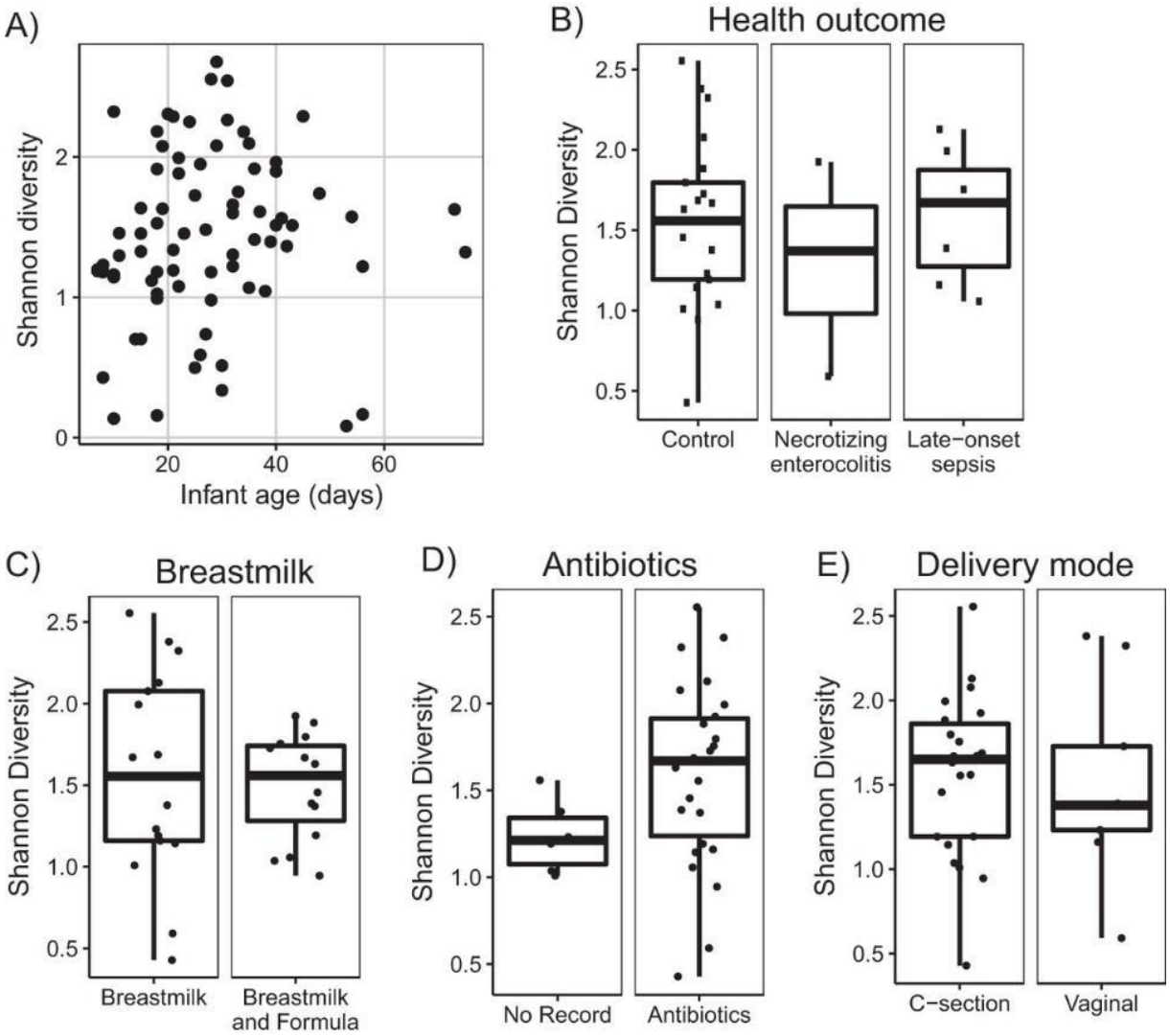


Figure 1.4 Alpha diversity measured by Shannon index of bacterial composition. A) Alpha diversity of all samples over the age of the infant. Boxplots of the average alpha diversity of each infant separated by B) health outcome, C) infants that were fed only breastmilk or a combination of formula and breastmilk, D) record of antibiotic usage, or E) delivery mode.

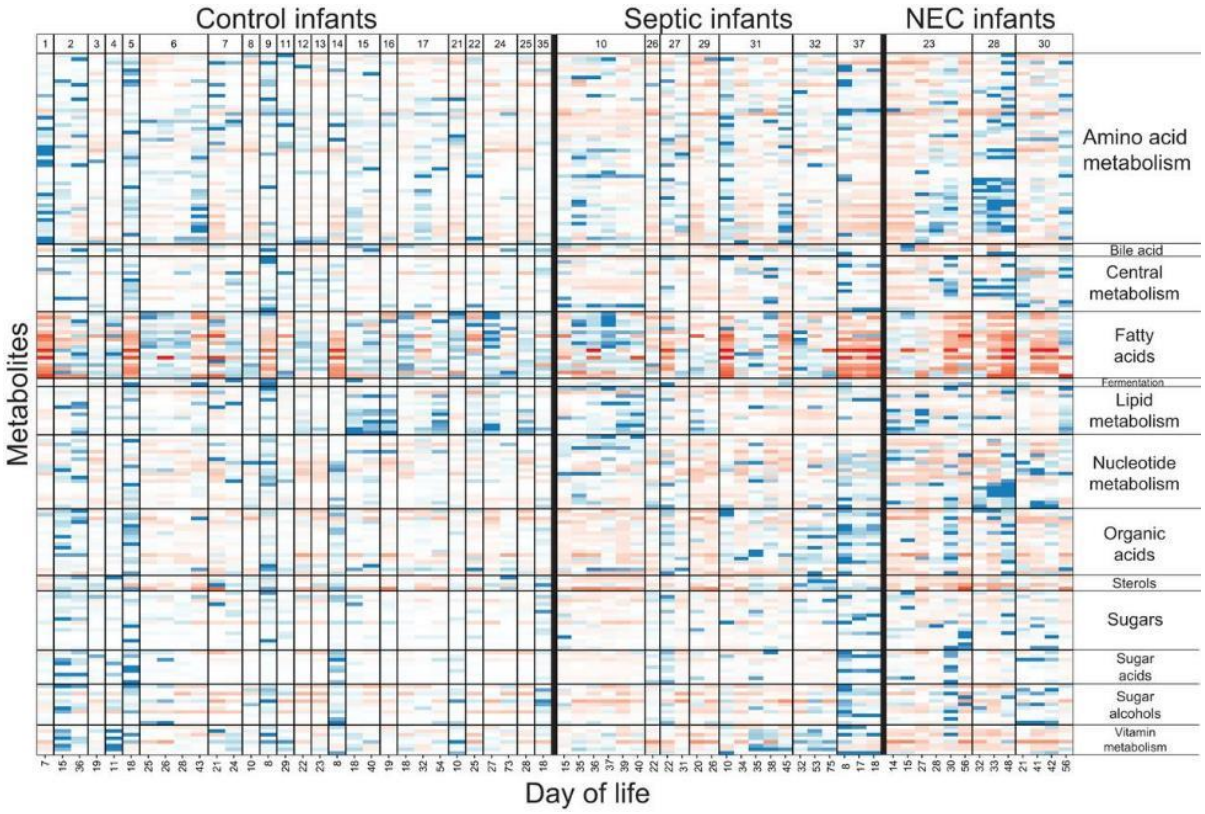


Figure 1.5 Metabolite profile of preterm infant fecal samples. Color is the modified z-score which is based on the median intensity for each metabolite in all infant samples. Red cells indicate standard deviations below the median and blue indicate standard deviations above the median value for each metabolite. Measured metabolites that could be assigned to a category are shown. Samples on the x-axis and grouped by infant and ordered longitudinally. Metabolites within each category are listed in the supplemental data.

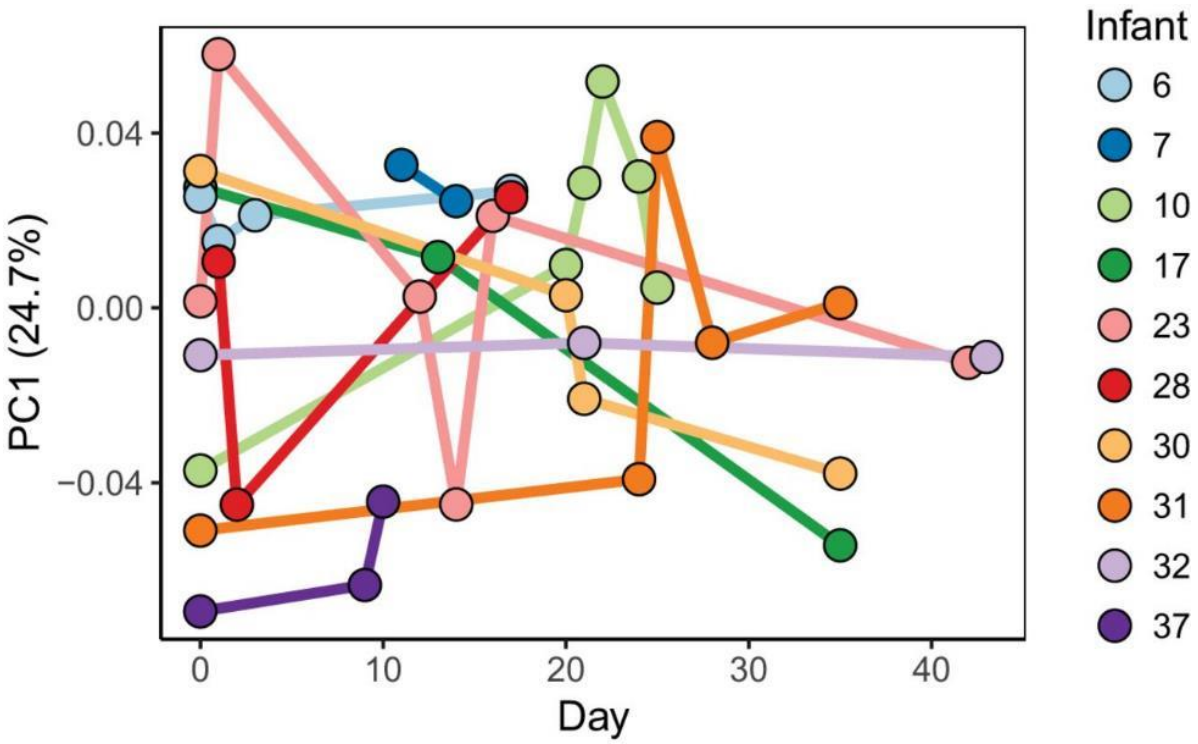


Figure 1.6 First component of PCoA of metabolite profile over time. Manhattan distances between samples were calculated and visualized by PCoA. The first principal component which explains the most variation among the samples is shown over time. Each dot represents a single fecal sample and is colored by infant. Lines connect samples for each infant to show change over time. Only infants with three or more longitudinal samples shown.

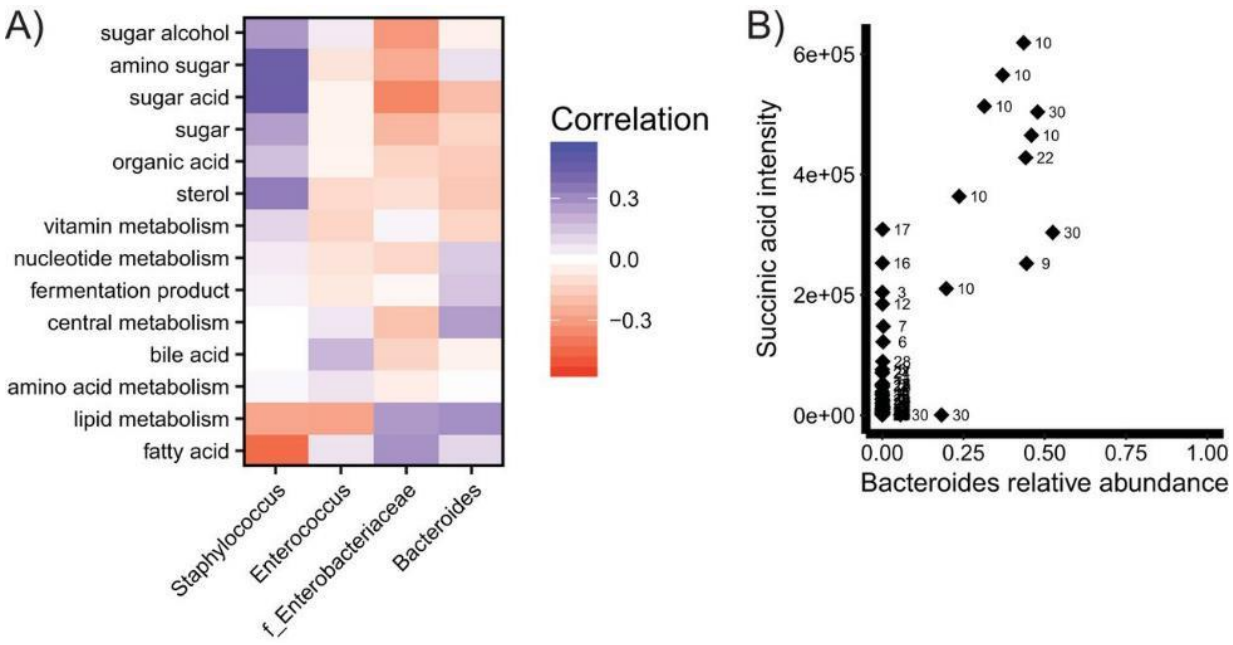


Figure 1.7 Correlations between bacteria abundances and metabolite intensities. A) Average of correlations between bacterial abundances and all metabolites in each metabolite category. B) Correlation between *Bacteroides* abundance and succinic acid intensity in all samples. Numbers indicate infant number.

Table 1.1. Clinical and sampling information for all infants.

Infant	**# of samples**	**Age sample(s) collected (days)**	**Age at disease onset (days)**	**Group**	**Fetal age at birth**	**Birth weight (g)**	**Antibiotics** (amp = ampicillin, gent=gentamycin)	**Delivery mode**	**Feeding type** (BM: breastmilk F: formula)	**Twin set**
1	2	7,7		control	27w4d	875		c-section	BM	
2	3	15,15,36		control	31w	1570	amp, gent	c-section	BM, F	
3	1	19		control	26w	980	amp, gent	c-section	BM	
4	2	11,11		control	30w3d	1335		vaginal	BM	
5	2	18,18		control	24w5d	630		c-section	BM	
6	4	25,26,28,43		control	28w5d	860	amp, gent	c-section	BM, F	
7	3	10,21,24		control	25w2d	885		c-section	BM, F	
8	1	10		control	25w4d	940	amp, gent	vaginal	BM	
9	1	8		control	27w2d	1205		vaginal	BM	
11	2	29,29		control	27w4d	850	amp, gent	vaginal	BM	
12	1	22		control	26w2d	880	amp, gent	c-section	BM, F	1
13	1	23		control	26w2d	925	amp, gent	c-section	BM, F	1
14	1	8		control	31w4d	1190	amp, gent	c-section	BM	
15	3	18,40,40		control	28w1d	1270	amp, gent	c-section	BM, F	2
16	1	19		control	28w1d	1355	amp, gent	c-section	BM, F	2
17	3	18,32,54		control	26w2d	660	amp, gent	c-section	BM	
21	1	10		control	28w6d	1180	amp, gent	c-section	BM	
22	1	25		control	28w6d	1360	amp, gent	vaginal	BM, F	
24	2	27,73		control	26w	740	amp, gent	c-section	BM	3
25	1	28		control	26w	780	amp, gent	c-section	BM	3
35	2	18,18		control	25w5d	920		c-section	BM, F	
23	7	14,15,27,28,30,30,56	27	NEC	26w6d	1080	amp, gent, cefotaxime, vancomycin	vaginal	BM	
28	4	31,32,33,48	31	NEC	26w	1060	vancomycin,piperacillin	c-section	BM, F	
30	4	21,41,42,56	41	NEC	23w6d	620	cefazolin,azithromycin,amp	vaginal	BM, F	
20	1	21	26	septic	24w5d	815	amp, gent	c-section	BM	
10	6	15,35,36,37,39,40	27	septic	26w5d	940	amp, gent, vancomycin	vaginal	BM, F	
26	1	22	22	septic	24w4d	660	amp, gent, cefotaxime, vancomycin	c-section	BM	4
27	2	22,31	29	septic	24w5d	650	amp, gent	c-section	BM	4
29	2	20,26	26	septic	26w1d	980	cefotaxime, vancomycin	c-section	BM	
31	5	10,34,35,38,45	34	septic	27w	710	amp, gent	c-section	BM, F	
32	4	32,32,53,75	32	septic	27w5d				BM, F	
37	3	8,17,18	13	septic	24w1d	680	amp, gent, cefazolin,oxacillin	vaginal	BM	

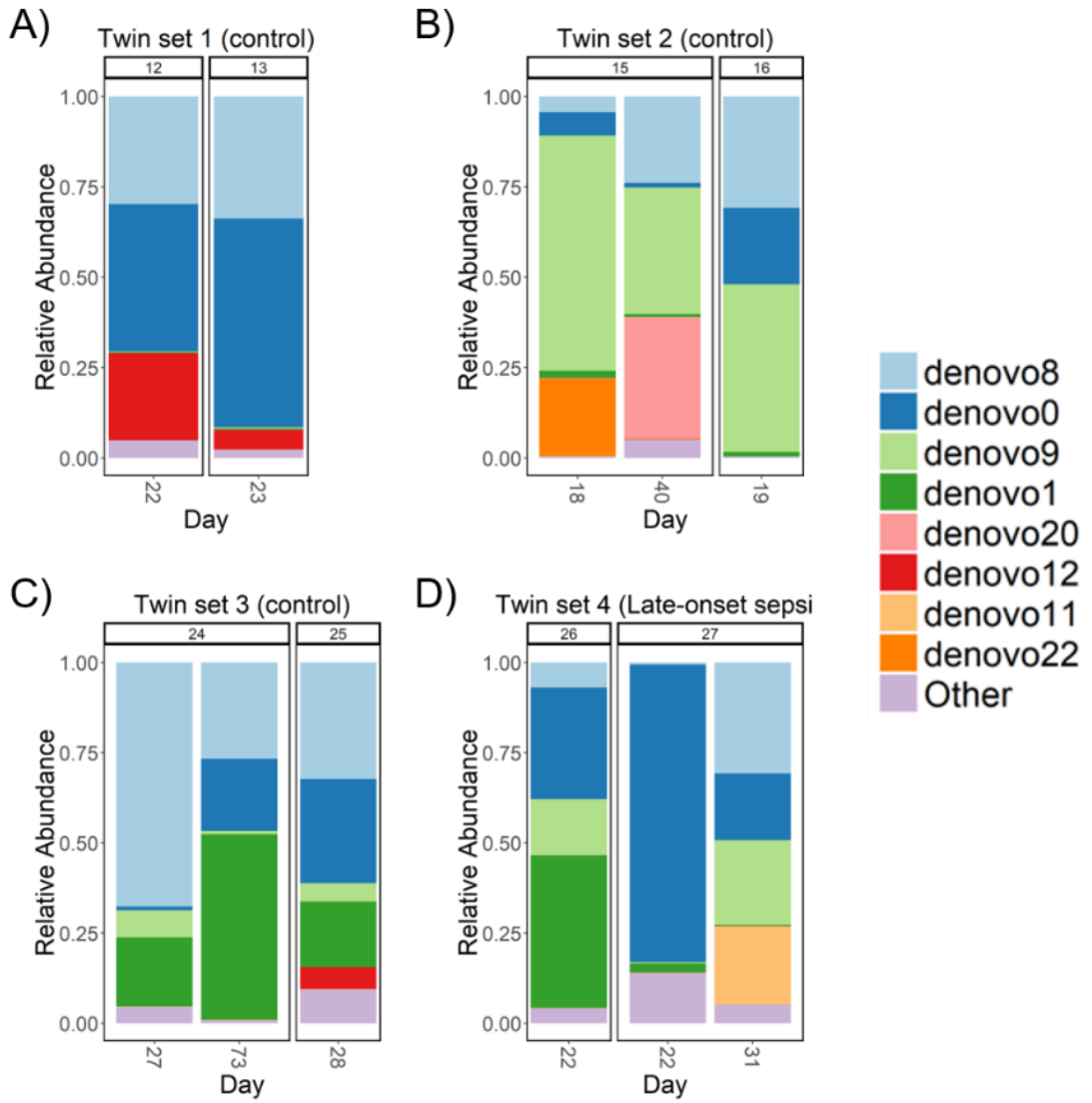


Figure S1.1 Bacterial composition of each set of twins.

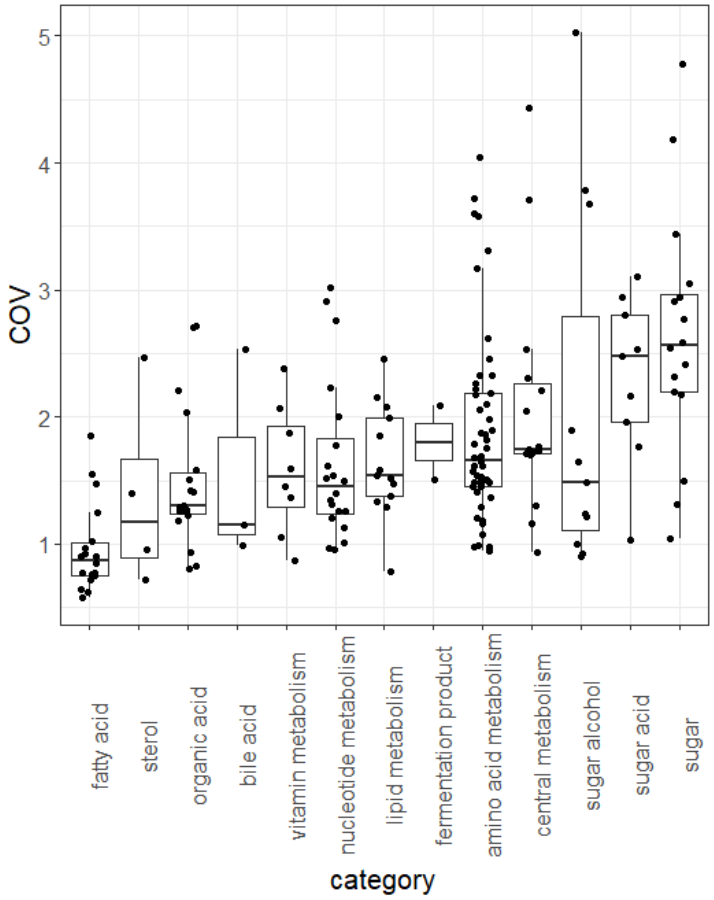


Figure S1.2 Average variation among infants of each metabolite grouped by category. Each dot represents a single metabolite. Coefficient of variation for each metabolite calculated as the standard deviation divided by the mean intensity of that metabolite in all samples from all infants.

Table S1.1 PERMANOVA for clinical factors explaining the differences in fecal a) bacterial composition and b) metabolome. Bolded factors are statistically significant. PERMANOVA for health outcome testes for differences among control, NEC, and late-onset sepsis. PERMANOVA for health and individual only includes infants with three or more longitudinal samples.

a)

	Df	**Variance explained (R^2)**	**P-value**
Health	2	0.079	0.64
Individual	**9**	**0.48**	**< 0.001**
Delivery mode	**1**	**0.12**	**0.046**
Breastmilk/Formula	1	0.07	0.076
Antibiotics	1	0.039	0.29
Date in NICU	6	0.25	.335

b)

	Df	**Variance explained (R^2)**	**P-value**
Health	2	0.15	0.095
Individual	**9**	**0.48**	**< 0.001**
Delivery mode	1	0.03	0.55
Breastmilk/Formula	**1**	**0.13**	**0.005**
Antibiotics	1	0.045	0.181

Table S1.2 All identified metabolites measured by GC-MS and their assigned categories.

Metabolite	Category
2-ketoisocaproic acid	amino acid metabolism
2-picolinic acid	amino acid metabolism
3,4-dihydroxyphenylacetic acid	amino acid metabolism
3-phenyllactic acid	amino acid metabolism
4-aminobutyric acid	amino acid metabolism
4-hydroxybenzoate	amino acid metabolism
4-hydroxyphenylacetic acid	amino acid metabolism
5-aminovaleric acid	amino acid metabolism
alanine	amino acid metabolism
alanine-alanine	amino acid metabolism
asparagine	amino acid metabolism
aspartic acid	amino acid metabolism
beta-alanine	amino acid metabolism
cysteine	amino acid metabolism
cystine	amino acid metabolism
glutamine	amino acid metabolism
glutaric acid	amino acid metabolism
glycine	amino acid metabolism
histidine	amino acid metabolism
homocystine	amino acid metabolism
homoserine	amino acid metabolism
indole-3-lactate	amino acid metabolism
isoleucine	amino acid metabolism
leucine	amino acid metabolism
lysine	amino acid metabolism
methionine	amino acid metabolism
methionine sulfoxide	amino acid metabolism
N-acetylaspartic acid	amino acid metabolism
n-acetylglutamate	amino acid metabolism
N-acetylornithine	amino acid metabolism
N-acetylputrescine	amino acid metabolism
N-methylalanine	amino acid metabolism
O-acetylserine	amino acid metabolism
ornithine	amino acid metabolism
phenol	amino acid metabolism
phenylacetic acid	amino acid metabolism
phenylalanine	amino acid metabolism
phenylethylamine	amino acid metabolism
pipecolinic acid	amino acid metabolism
proline	amino acid metabolism
putrescine	amino acid metabolism
pyrrole-2-carboxylic acid	amino acid metabolism
serine	amino acid metabolism
shikimic acid	amino acid metabolism
threonine	amino acid metabolism
trans-4-hydroxyproline	amino acid metabolism
tryptophan	amino acid metabolism
tyramine	amino acid metabolism
tyrosine	amino acid metabolism
urea	amino acid metabolism
urocanic acid	amino acid metabolism
valine	amino acid metabolism
N-acetyl-D-galactosamine	amino sugar
n-acetyl-d-hexosamine	amino sugar
chenodeoxycholic acid	bile acid
cholic acid	bile acid
taurine	bile acid
aconitic acid	central metabolism
alpha-ketoglutarate	central metabolism
citramalic acid	central metabolism
citric acid	central metabolism
dihydroxyacetone	central metabolism
fumaric acid	central metabolism
glucose-6-phosphate	central metabolism
glyceric acid	central metabolism
isocitric acid	central metabolism

malic acid	central metabolism
oxalic acid	central metabolism
pyrophosphate	central metabolism
pyruvic acid	central metabolism
succinic acid	central metabolism
sulfuric acid	central metabolism
2-hydroxybutanoic acid	fatty acid
arachidic acid	fatty acid
arachidonic acid	fatty acid
behenic acid	fatty acid
capric acid	fatty acid
cerotinic aci	fatty acid
elaidic acid	fatty acid
heptadecanoic acid	fatty acid
isoheptadecanoic acid NIST	fatty acid
lauric acid	fatty acid
lignoceric acid	fatty acid
linoleic acid	fatty acid
myristic acid	fatty acid
nonadecanoic acid	fatty acid
oleic acid	fatty acid
palmitic acid	fatty acid
pentadecanoic acid	fatty acid
stearic acid	fatty acid
butane-2,3-diol NIST	fermentation product
lactic acid	fermentation product
1-monoheptadecanoyl glyceride NIST	lipid metabolism
1-monoolein	lipid metabolism
1-monopalmitin	lipid metabolism
1-monostearin	lipid metabolism
2-monoolein	lipid metabolism
2-monopalmitin	lipid metabolism
dodecanol	lipid metabolism
ethanolamine	lipid metabolism
glycerol-alpha-phosphate	lipid metabolism
monomyristin	lipid metabolism
phosphoethanolamine	lipid metabolism
phytol	lipid metabolism
propane-1,3-diol NIST	lipid metabolism
3-aminoisobutyric acid	nucleotide metabolism
3-hydroxypropionic acid	nucleotide metabolism
5,6-dihydrouracil	nucleotide metabolism
7-methylguanine NIST	nucleotide metabolism
adenine	nucleotide metabolism
adenosine	nucleotide metabolism
adenosine-5-monophosphate	nucleotide metabolism
cytosin	nucleotide metabolism
guanosine	nucleotide metabolism
inosine	nucleotide metabolism
orotic acid	nucleotide metabolism
pseudo uridine	nucleotide metabolism
ribose	nucleotide metabolism
thymidine	nucleotide metabolism
thymine	nucleotide metabolism
uracil	nucleotide metabolism
uric acid	nucleotide metabolism
uridine	nucleotide metabolism
xanthine	nucleotide metabolism
xanthosine	nucleotide metabolism
2,3-dihydroxybutanoic acid NIST	organic acid
2-deoxytetronic acid	organic acid
2-deoxytetronic acid NIST	organic acid
2-hydroxyglutaric acid	organic acid
2-hydroxyhexanoic acid	organic acid
2-hydroxyvaleric acid	organic acid
3-hydroxy-3-methylglutaric acid	organic acid
3-hydroxybutyric acid	organic acid
adipic acid	organic acid
azelaic acid	organic acid
benzoic acid	organic acid
digalacturonic acid	organic acid
erythronic acid lactone	organic acid
glycolic acid	organic acid
hexaric acid	organic acid
maleic acid	organic acid
syringic acid	organic acid
tartaric acid	organic acid
cholesterol	sterol
dihydrocholesterol	sterol

squalene	sterol
stigmasterol	sterol
beta-gentiobiose	sugar
cellobiose	sugar
fructose	sugar
glucose	sugar
glycerol-3-galactoside	sugar
levoglucosan	sugar
lyxose	sugar
maltose	sugar
maltotriose	sugar
mannose	sugar
N-acetyl-D-mannosamine	sugar
raffinose	sugar
tagatose	sugar
trehalose	sugar
xylose	sugar
xylulose NIST	sugar
galacturonic acid	sugar acid
gluconic acid	sugar acid
gluconic acid lactone	sugar acid
lactobionic acid	sugar acid
ribonic acid	sugar acid
saccharic acid	sugar acid
threonic acid	sugar acid
UDP-glucuronic acid	sugar acid
xylonic acid	sugar acid
erythritol	sugar alcohol
erythrose	sugar alcohol
fucose	sugar alcohol
galactinol	sugar alcohol
hexitol	sugar alcohol
isothreitol	sugar alcohol
lactitol	sugar alcohol
lyxitol	sugar alcohol
mannitol	sugar alcohol
palatinitol	sugar alcohol
xylitol	sugar alcohol
4-pyridoxic acid	vitamin metabolism
delta-tocopherol NIST	vitamin metabolism
hexuronic acid1	vitamin metabolism
nicotinamide	vitamin metabolism
nicotinic acid	vitamin metabolism
tocopherol alpha-	vitamin metabolism
tocopherol beta NIST	vitamin metabolism
tocopherol gamma-	vitamin metabolism
1,2-anhydro-myo-inositol NIST	xother
2-deoxypentitol NIST	xother
2-hydroxypyrazinyl-2-propenoic acid ethyl ester NIST	xother
4-hydroxymandelic acid	xother
5-hydroxymethyl-2-furoic acid NIST	xother
6-hydroxynicotinic acid	xother
acetophenone NIST	xother
allantoic acid	xother
alloxanoic acid NIST	xother
aminomalonate	xother
beta-mannosylglycerate	xother
butyrolactam NIST	xother
caffeine	xother
conduritol-beta-epoxide	xother
creatinine	xother
docosahexaenoic acid	xother
epsilon-caprolactam	xother
hexadecylglycerol NIST	xother
hydroxycarbamate NIST	xother
hydroxylamine	xother
indole-3-acetate	xother
inositol-4-monophosphate	xother
isohexonic acid	xother
isothreonic acid	xother
maleimide	xother
malonic acid	xother
methyl O-D-galactopyranoside	xother
methyltetrahydrophenanthrenone 1 NIST	xother
parabanic acid NIST	xother
sebacic acid, di(2-octyl) ester NIST	xother
tyrosol	xother
xylonolactone NIST	xother
zymosterol	xother

CHAPTER 2

Making it last: Storage time and temperature have differential impacts on metabolite profiles of airway samples from cystic fibrosis patients

Co-authors: Lisa Carmody, Tara Gallagher, John LiPuma, and Katrine Whiteson

ABSTRACT

Metabolites of human or microbial origin have the potential to be important biomarkers of disease state in cystic fibrosis (CF). Clinical sample collection and storage conditions may impact metabolite abundances with clinical relevance. We measured the change in metabolite composition based on untargeted gas chromatography mass spectrometry (GC-MS) when CF sputum samples were stored at either 4°C, -20°C, or -80°C with one or two freeze-thaw cycles. Daily time points were taken for one week and then weekly for 4 weeks (4°C) and 8 weeks (-20°C). The metabolites in samples stored at -20°C maintained similar abundances compared to -80°C over the course of eight weeks (average change in Bray-Curtis distance: 0.06±0.04) and were also stable after one or two freeze-thaw cycles. However, metabolite profiles of samples stored at 4°C shifted after one day and continued to change over the course of four weeks (average change in Bray-Curtis distance: 0.31±0.12). Several amino acids and other metabolite abundances increased with time when stored at 4°C, but remained constant at -20°C. Storage temperature was a significant factor driving the metabolite composition (PERMANOVA R^2 = 0.32 to 0.49, p= <0.001). CF sputum samples stored at -20°C at the time of sampling maintain a relatively stable untargeted GC-MS profile. Samples should be frozen on the day of collection, as more than one day at 4°C impacts the global composition of the metabolites in the sample.

IMPORTANCE Metabolomics has great potential for uncovering biomarkers of disease state in CF and many other contexts. However, sample storage timing and temperature may alter the abundance of clinically relevant metabolites. In order to assess whether existing samples are stable and to direct future study design, we conducted untargeted GC-MS metabolomics on CF sputum samples after one or two freeze-thaws and storage at 4°C and -20°C for four to eight weeks. Overall, storage at -20°C and freeze-thaw cycles had little impact on metabolite profiles; however, storage at 4°C shifted metabolite abundances significantly. GC-MS profiling will aid in our understanding of the CF lung, but care should be taken in studies using sputum samples to ensure samples are properly stored.

INTRODUCTION

The staggering metabolic complexity of any human biologic specimen results from both human and microbial metabolism, and could provide clinically relevant biomarkers of disease. Recent studies based on chromatography and mass spectroscopy estimate that tens or hundreds of thousands of distinct metabolites are in human biologic specimens, and that a third to half of them are produced or altered by microbes (Dorrestein, Mazmanian, and Knight 2014; Wikoff et al. 2009). In this study, we assessed the impact of storage temperature and time on metabolite composition in sputum samples collected from people with cystic fibrosis (CF). CF is a genetic disease caused by a mutation in a cellular ion transporter that results in increased susceptibility of the respiratory tract to bacterial infection (P M Quinton 1983; Paul M Quinton 2008; Stoltz, Meyerholz, and Welsh 2015). The lives of persons with CF are punctuated by periods of worsened respiratory symptoms referred to as pulmonary exacerbations. Although the etiology of these events remains unclear, changes in the structure and/or activity of airway microbial communities is

believed to play a role (Carmody et al. 2015; Jiangchao Zhao et al. 2012; Laguna et al. 2015; Lynch and Bruce 2013; D. J. Smith et al. 2014; Stressmann et al. 2011; Twomey et al. 2013). Decades of study of the microbiology of CF airway infection have relied on recovery of select bacterial species in culture of respiratory specimens. More recently, next-generation DNA sequencing and metabolomic analyses have contributed to a broader appreciation of the complexity and dynamics of the microbial ecology of CF airways (Price et al. 2013; Carmody et al. 2015; Jiangchao Zhao et al. 2012). Culture independent approaches have potential to identify changes in microbial community structure and activity associated with pulmonary exacerbations, thereby offering insight to novel approaches to prevent or better treat these episodes (Robroeks et al. 2010; Bos, Sterk, and Schultz 2013; Bos et al. 2014, 2016).

The utility of metabolomics in advancing our understanding of CF exacerbations depends, however, on an appreciation of how variables in sample handling may impact measures of global metabolic profile and/or assessment of specific metabolites of interest. Respiratory samples, including expectorated sputum and bronchoalveolar lavage (BAL) fluid, are most often obtained from patients in clinical settings where immediate freezing and storage is not practical. Samples may be held at 4°C or -20˚C for variable periods of time before being stored at -80˚C. Samples may also undergo repeated cycles of freezing and thawing prior to analysis. While studies have addressed the impact of storage temperature on metabolomic analysis of clinical samples such as plasma and urine (Krug et al. 2012; Rist et al. 2013; Pinto et al. 2014), comparable studies assessing sputum are lacking. We previously examined the impact of storage conditions on CF sputum metabolomic profiles based on liquid chromatography mass spectrometry (LC-MS) and

found that profiles were stable in samples stored for at least four weeks at -20˚C (Jiangchao Zhao et al. 2015). While LC-MS can be used to study low-molecular weight compounds, it cannot ionize nonpolar compounds and is not practical for analysis of volatile compounds, which are better analyzed using gas chromatography mass spectrometry (GC-MS). With respect to CF, recent breath sampling has shown that levels of the volatile 2,3-butanedione, a pH-neutral bacterial fermentation product, were elevated in the breath of CF patients compared to healthy individuals. Another molecule from the same metabolic pathway, 2,3-butanediol, was associated with increased virulence of the CF opportunistic pathogen *Pseudomonas aeruginosa* in both culture based and murine infection models (Venkataraman et al. 2011, 2014; Whiteson et al. 2014). These studies highlight the potential of metabolomic analysis in investigating the microbiology of airway infection in CF. A critical element in such studies, however, is a more complete understanding of the impact of sample storage on measures of metabolites detected by GC-MS. In the study reported here, we assessed the stability of a wide range of volatile metabolites in sputum samples stored at 4˚C and -20˚C for variable periods of time with one or two freeze-thaw cycles.

RESULTS

Bacterial community profiles

Bacterial 16S rRNA gene sequencing revealed similar profiles for two sputum samples collected from two CF patients during the course of routine medical care (**Figure 2.1**). *Prevotella* and *Pseudomonas* dominate in both samples; the distribution of microbes in these samples is typical of an adult CF airway community.

Effect of storage at 4°C or -20°C

Our goal was to determine the impact of storage time and temperature on GC-MS metabolite profiles, as depicted in the flowchart in **Figure 2.2**. Two sputum samples were split into 76 aliquots and stored at 4°C or -20°C, with 2 additional aliquots stored immediately at -80°C; daily time points were taken for a week and weekly time points for 4 weeks (4°C) and 8 weeks (-20°C) (**Figure 2.2**). GC-MS was performed on each sample and the resulting metabolomic profiles were visualized by principal coordinate analysis (PCoA) (**Figure** 2.**3**). A total of 239 metabolites were detected, of which 104 metabolites were identifiable. A subset of molecules are represented in a heat map depicting the fold change difference in metabolite intensities at each storage temperature over time (**Figure S2.1, Figure S2.2**). For both patients, the samples stored at -20°C clustered together with their respective time 0 samples that were stored immediately at -80°C, while samples stored at 4°C drifted in the ordination space over time. Storage temperature at 4°C versus -20°C was found to significantly affect the overall metabolite profile (PERMANOVA r^2= 0.49 (sample 1); r^2= 0.32 (sample 2), $p < 0.001$) (Supplemental Table 1). Storage at 4˚C resulted in larger changes in metabolite abundances over time (**Figure S2.3**). A major contributor to the changing metabolite profile at 4°C was the increasing total ion count between days 1 and 28, which remained stable at -20°C (**Figure S2.4**). The Bray Curtis distance of each metabolomic profile from the time 0 profile was calculated for samples stored at 4°C or -20°C (**Figure 2.4**). After one day, the metabolomic profiles for samples stored at 4°C were observed to be more distant than samples stored at 20°C, and the average distance from time 0 increased each week for samples stored at 4°C. The average distance of the metabolomic profiles of the samples stored at -20°C from the time 0 profile remained steady over the course of 8 weeks (average change in Bray-Curtis distance: 0.06±0.04),

indicating that the overall metabolite profile did not change when stored at -20°C. In contrast, metabolite profiles of samples stored at 4°C continued to change over the course of four weeks (average change in Bray-Curtis distance: 0.31±0.12).

Metabolites distinguishing patient samples

The samples collected from each patient clustered separately along the primary PCoA axis (**Figure 2.3**), indicating that the two patients' sputa have unique metabolomics profiles. The metabolites that differed the most between the samples from patient 1 and patient 2 as determined by supervised Random Forest analysis and ranked by mean decrease in predictive accuracy were putrescine, xylitol, glycerol, 5-aminovaleric acid, and uric acid (**Figure 2.5A**).

Metabolites that change with storage temperature and time

The metabolite profiles of sputum samples from both patients stored at 4°C for varying lengths of time separate along the secondary PCoA axis, indicating common changes in overall metabolomic profile with storage at this temperature (**Figure 2.3**). A supervised Random Forest analysis was used to determine which metabolites were responsible for the differences between samples stored at 4°C and -20°C. The top variables of importance (VIPs) separating samples collected at different temperatures (**Figure 2.5B**). Many of the metabolites that are most different at each storage temperature are amino acids (**Figure 2.5B, 2.5C**). The intensities of the VIP metabolites at each time point as shown in **Figure 2.5C** and indicate that the amino acids are increasing in abundance with time at 4°C while staying more constant at -20°C.

Clear trends showing either an increase or decrease over time at either 4°C or -20°C were not observed for several metabolites that have been determined to be clinically

relevant to CF in other studies, including lactic acid, putrescine, and 5-aminovaleric acid (**Figure 2.6**).

Impact of freeze-thaw on metabolomics profiles

At each time point, samples were assessed after thawing once, and again after a second freeze-thaw. The differences in overall metabolite profiles of samples thawed once or twice were not statistically significant (PERMANOVA $r^2 = 0.005$, $p = 0.72$, see Supplemental Table 1).

Metabolite intensity and the coefficient of variance

The coefficient of variation (COV: standard deviation / mean) for each metabolite stored at either temperature was assessed (**Figure 2.7**). The metabolites in aliquots stored at -20°C had lower COVs compared to those at 4°C as demonstrated by the distribution of points in the violin plot in **Figure 2.7**, where the aliquots stored at -20°C largely fall under 0.3 while the COVs for the aliquots stored at 4°C are more widely distributed. The distribution of COVs for each metabolite during storage at -20°C remain about the same regardless of whether the samples at -20°C were stored for 28 or 56 days, or whether they were thawed once or twice (data not shown).

DISCUSSION

Metabolite profiling complements other culture-independent approaches such as bacterial 16S rRNA gene amplicon sequencing in characterizing bacterial communities in biologic specimens. This could aid in identifying biomarkers of disease progression that are believed to be associated with changes in the activity of host microbial communities. In CF, for example, production of fermentation products such as lactic acid have been found to be an indicator of pulmonary exacerbation that can be measured with GC-MS (Zang et al.

2017; Tate et al. 2002; Twomey et al. 2013). Sample storage conditions can have a significant impact on the abundances of metabolites in biologic samples, and studies assessing metabolite profiles need to account for these effects. In previous work, we observed that CF sputum sample metabolite profiles determined by LC-MS were unstable after storage at 4°C (Carmody et al. 2015).

In this study, we captured smaller, aqueous and more volatile metabolites in CF sputum with an acetonitrile:IPA:water extraction followed by untargeted GC-MS. Unfortunately, some microbial fermentation products of interest, such as 2,3-butanedione and 2,3-butanediol, are extremely volatile and were not consistently detectable with this approach. Profiles of two sputum samples stored at either 4°C or -20°C for varying lengths of time were assessed. We observed that storage at -20°C yielded stable metabolite profiles with similar variation in metabolite abundances as found between technical replicates. More specifically, the range of Bray-Curtis dissimilarity values among samples stored at -80°C did not increase among samples stored at -20°C. In contrast, the Bray-Curtis distances of samples stored at 4°C increased significantly with storage time. Of note, however, some metabolites appeared to remain stable during storage at 4°C, including some with potential as biomarkers of CF clinical disease state, such as lactic acid.

Inter-individual variation is often the most significant factor driving differences in microbial or metabolite composition between biological specimens, although the relative contributions of human and microbial metabolism to metabolite profiles are often difficult to distinguish. In this study, the metabolite profiles of samples from each of two individuals were consistently distinct, with the unique features of each resulting in clear demarcation by PCoA. Sample storage temperature also had a significant impact on metabolite profiles

that were also apparent by PCoA. The impact of sample storage time and temperature were apparent in analysis of total metabolite intensity versus time (**Figure S2.1);** the total abundance of the metabolites being measured increased at 4°C, but not at -20°C. Overall, these results suggest either active microbial metabolism or metabolite degradation in cells at 4°C, leading to increases in the intensity of compounds detectable by GC including amino acids and other cellular debris.

It is also possible that changes in bacterial community structure under different storage conditions account for differences in metabolite abundances. Previous work has shown, however, that differences in community structure between sputum sample aliquots stored at room temperature and aliquots stored at other temperatures were less than differences observed between intra- and inter-sample controls (J Zhao et al. 2011). In samples with communities not yet dominated by *Pseudomonas aeruginosa* a reduction in diversity after storage at room temperature was observed (J Zhao et al. 2011).

We note that although this study involved GC-MS analysis of nearly 200 samples, it was limited to two 'parent' sputum samples from which dozens of aliquots were obtained, variably stored and analyzed. While we have no reason to believe that these samples were atypical in any regard, previous work has shown that persons with CF harbor distinct microbial communities, at least during the early and intermediary stages of lung disease (Carmody et al. 2015; Jiangchao Zhao et al. 2012). An analysis of metabolite stability (during various storage conditions) in sputum samples from a greater number of individuals will be required to better demonstrate the generalizability of our findings.

In summary, storage of CF sputum samples at 4°C leads to changes in metabolite profiles within a day, with greater variation in metabolite abundances and an increase in

the abundance of many of the metabolites, including several amino acids detected by untargeted GC-MS profiling. Nevertheless, several metabolites of clinical interest remain stable, including lactic acid, putrescine and 5-aminovaleric acid. Our results suggest that CF sputum samples stored at -20°C retain stable GC-MS profiles for as long as two months. Freezing and thawing samples once or twice does not have a significant effect on metabolite intensities. These findings need to be considered in designing studies to assess the metabolome of microbial communities in CF airways and other environments.

MATERIALS AND METHODS

Sputum collection, storage and sequencing

CF sputum samples were collected spontaneously from two patients during the course of routine medical care. Sample collection was approved by the University of Michigan Institutional Review Board. Samples were kept on ice for up to 30 min prior to processing. An equal volume of cold sterile water was added to each sample before mechanical homogenization with a sterile tissue homogenizer (Omni International) on ice for 10 seconds. Each sample was divided into 100µL aliquots and duplicate or quadruplicate aliquots were stored at different temperatures for various lengths of time before being stored at -80°C (**Figure 2**). At each time point, two aliquots stored at -20°C were thawed on ice for 30 min before being stored at -80°C.

The bacterial communities in each of the two sputum samples (stored continuously at

-80°C) were characterized as described previously (Carmody et al. 2015). In brief, the V4 region of the 16S rRNA gene was amplified, sequenced, and analyzed with mothur.

Metabolite extraction and metabolomics profiling

A total of 156 sputum samples were shipped frozen on dry ice to the West Coast Metabolomics Center at UC Davis for untargeted GC-TOF profiling. Metabolites were extracted from 20 microliters of sputum with 1mL of 3:3:2 acetonitrile:isopropyl alcohol:water before derivatization and analysis by GC-MS using Fiehnlab Standard operating procedures (Cajka and Fiehn 2016; Kind et al. 2009). Metabolites were identified by comparison to the binbase database. Metabolite intensities were normalized to the total intensity of identified metabolites for each sample. Intensities were reported for 239 metabolites, of which 104 could be identified.

Statistical analysis

Principal coordinate analysis (PCoA) based on Bray-Curtis (BC) distances was used to visualize differences in the metabolite profiles between samples and processing conditions. All analyses were done with R (3.2.5). PCoA was performed using the ape library. BC distance was calculated using the vegdist function in vegan library. PERMANOVA was performed using the Adonis test in the vegan package in R (Oksanen et al. 2015). Random Forest analysis was carried out with the randomForest package in R with default parameters (ntree = 500) (Liaw and Wiener 2002). Plots were made using the ggplot2 and ggpubr libraries.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the help of several people from the West Coast Metabolomics Center at UC Davis including Megan Showalter, Sili Fan and Professor Oliver Fiehn.

FUNDING INFORMATION

This work was supported by grants from the National Heart, Blood, Lung Institute

(R56HL126754-01A1) to JJL and KW, and the Cystic Fibrosis Foundation (to JJL). KW is supported by a Gilead CF Research Scholars award (app_00b072) and a UC Davis West Coast Metabolomics Center pilot grant (DK097154), where metabolomics data were collected.

FIGURES AND TABLES

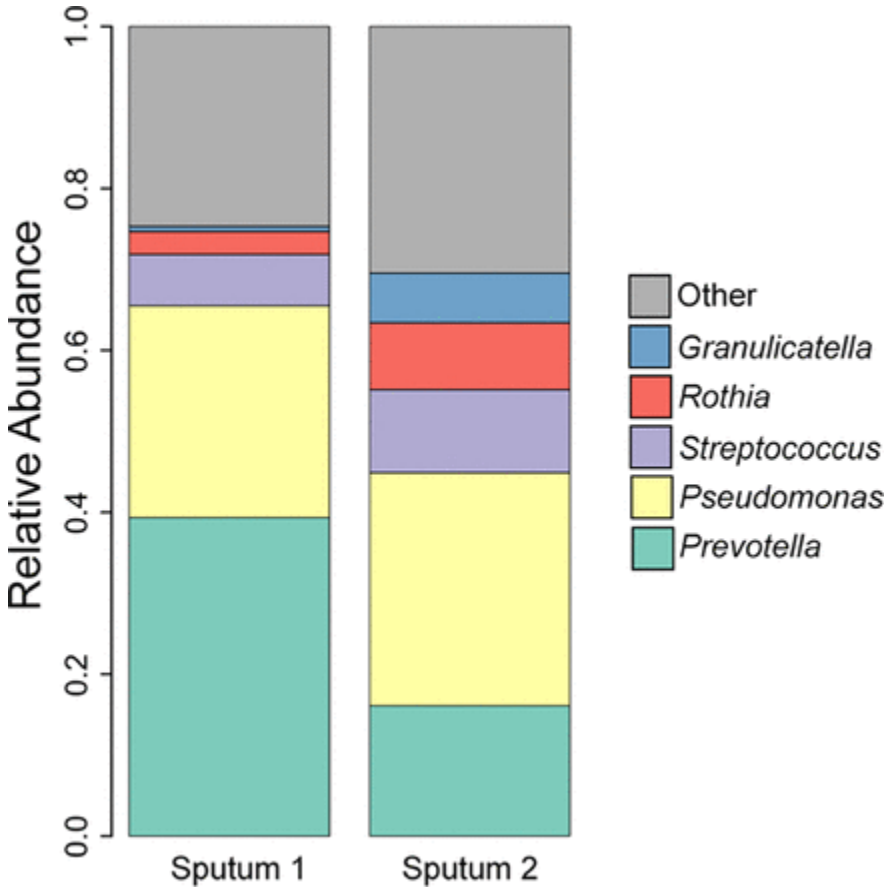
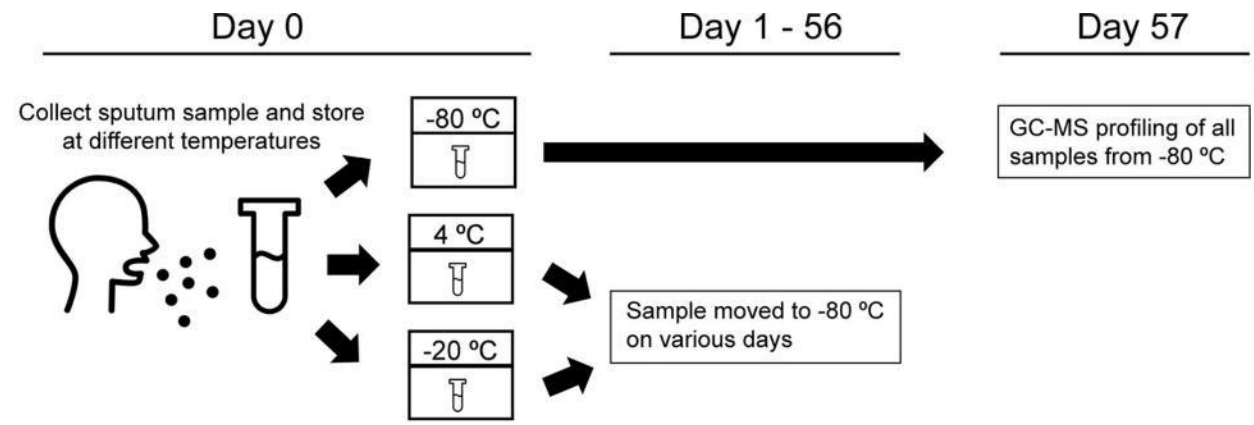


Figure 2.1. 16S rRNA gene sequence profiles of bacterial communities for each of the sputum samples that were aliquoted for the storage study. The V4 region of the 16S rRNA gene was amplified, sequenced, and analyzed with mothur as described previously (Jiangchao Zhao et al. 2015). Relative abundances of the five most abundant taxa in each sputum sample are shown.



Days with metabolite storage data

	1	2	3	4	5	6	7	14	21	28	35	42	49	56
-20 ºC	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4 ºC	X	X	X	X	X	X	X	X	X	X				

Figure 2.2. Flow chart and study design schematic. Sputum samples from two patients were homogenized, aliquoted, and stored at either -20°C or 4°C. Duplicates of stored samples were taken daily for a week, and weekly samples were taken for 4 weeks at 4°C and 8 weeks at -20°C. At each time point, an additional replicate was subjected to either one or two freeze thaw cycles. The aliquots were all analyzed with untargeted Gas Chromatography – Mass Spectroscopy (GC-MS).

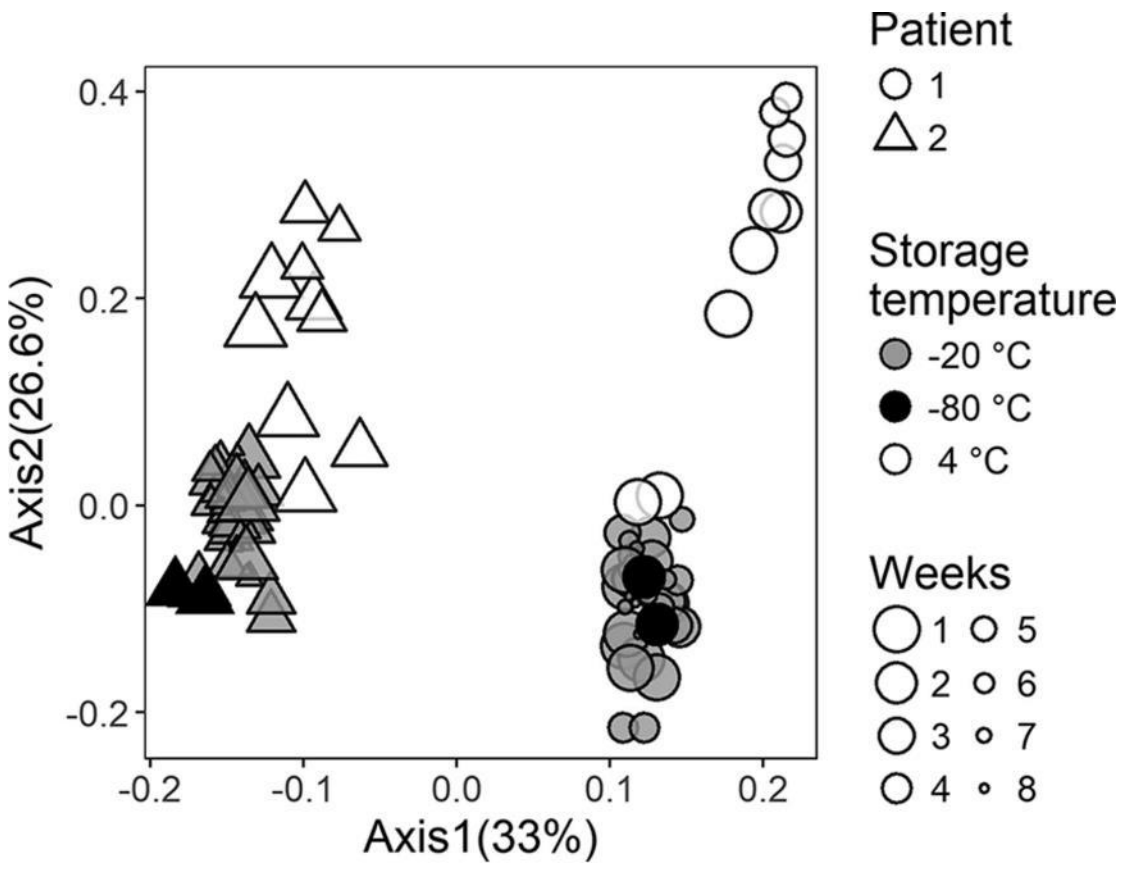


Figure 2.3. Principal coordinates analysis (PCoA) of Bray-Curtis distances calculated from the metabolite abundances in each sample. The percent variance explained by each axis is shown in parenthesis in the axis label. Aliquots of the sample from Patient 1 are shown in circles while those from the sample from Patient 2 are depicted by triangles. The gold standard -80°C sample are represented by solid black circles, while -20°C samples are grey and 4°C samples are open circles. The size of the symbol reflects storage time with later times represented as smaller symbols.

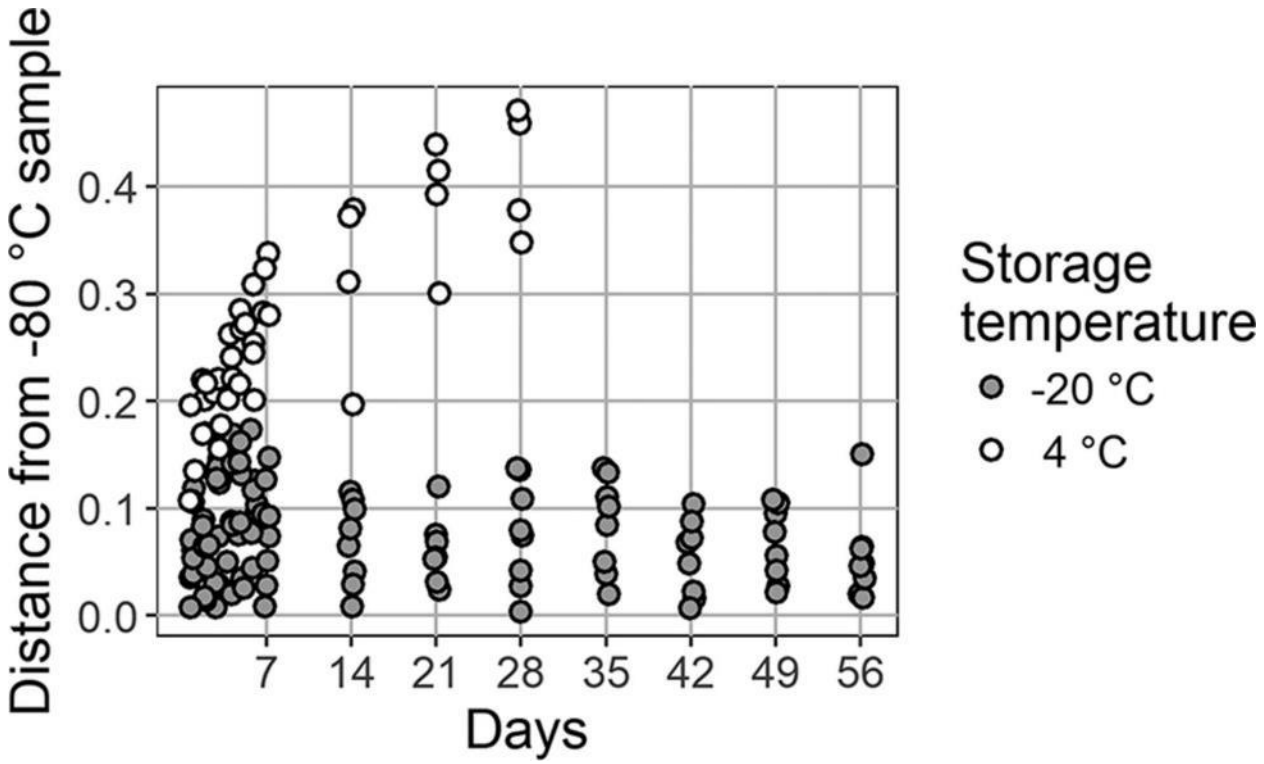


Figure 2.4. Effect of storage at -20°C versus 4°C. Bray-Curtis distances between samples at each storage temperature and the corresponding patient samples stored immediately at -80°C are shown. All replicates from both samples are shown, including those subjected to one or two freeze thaw cycles.

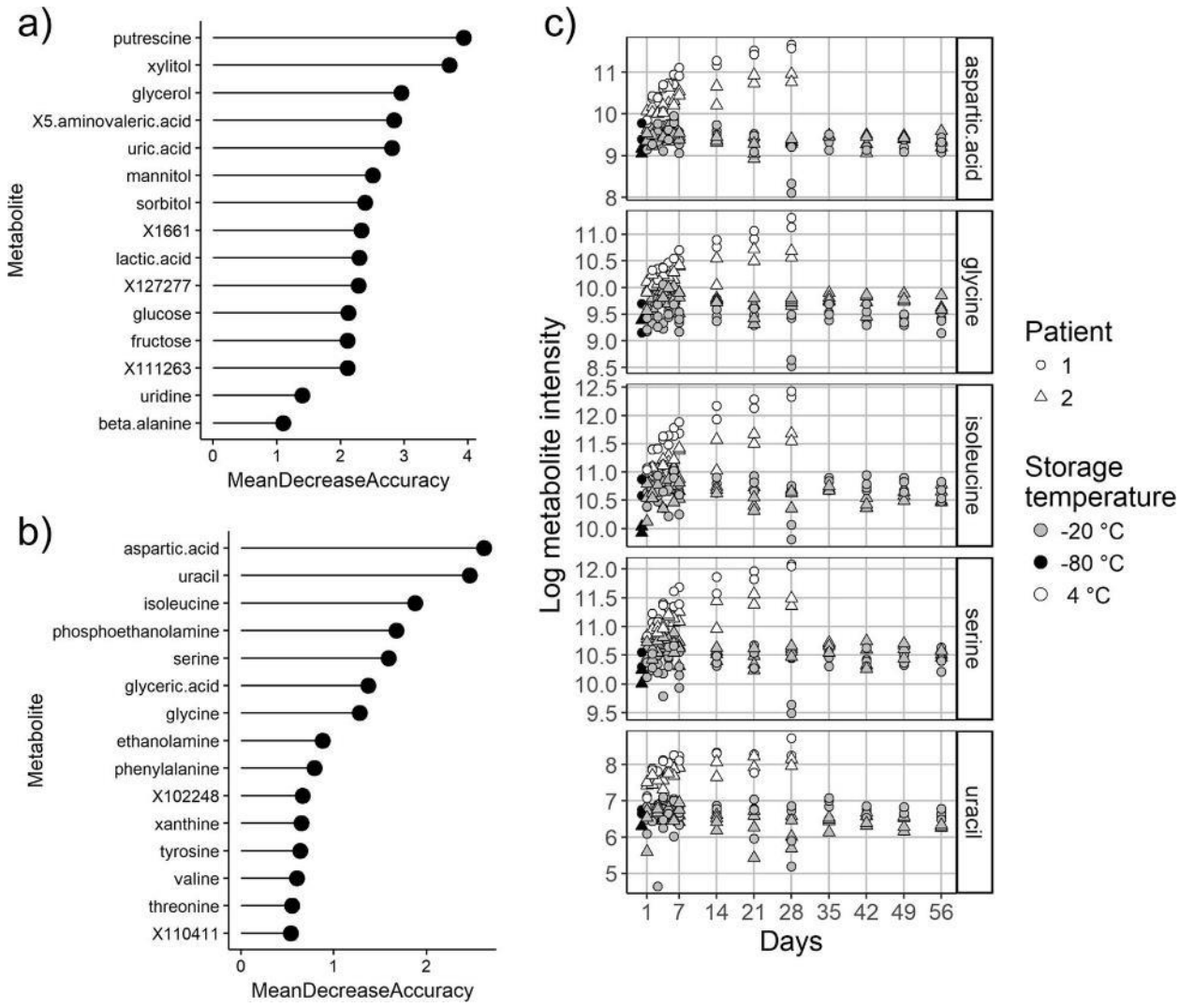


Figure 2.5. The variables of importance from a supervised Random Forest analysis to determine which metabolites best distinguish (A) patient 1 and patient 2 samples and (B) storage at -20°C vs 4°C. All time points were included and intensities were log transformed. (C) Intensities over time of five metabolites that distinguish storage temperature shown.

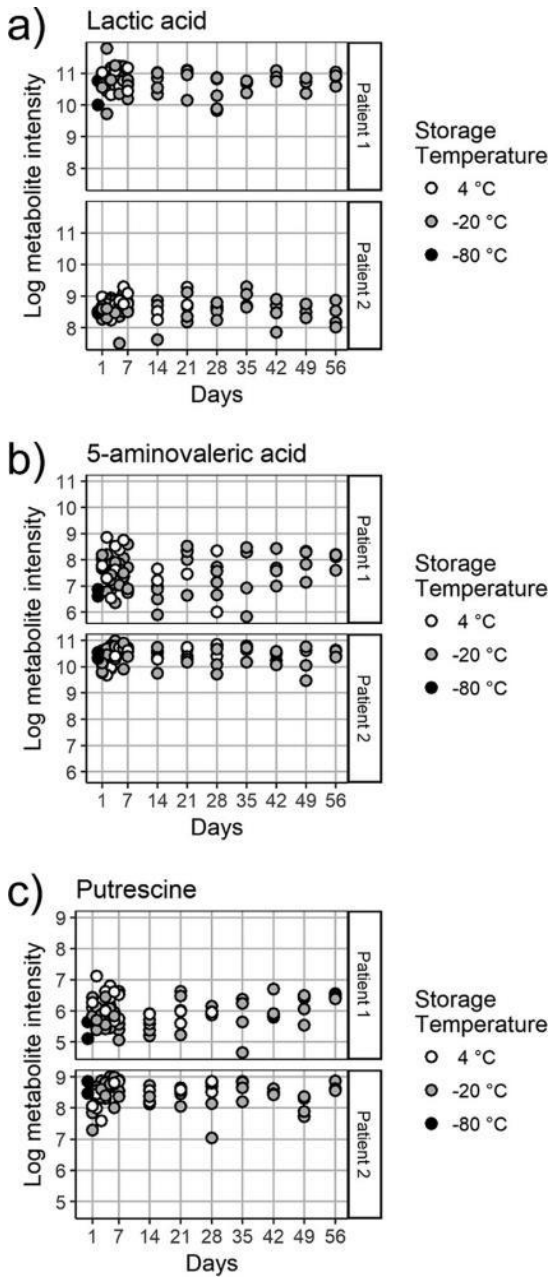


Figure 2.6. Intensities of several metabolites of potential clinical significance after storage at 4°C or -20°C for patients 1 and 2, including (A) lactic acid, (B) 5-aminovaleric acid, and (C) the polyamine putrescine.

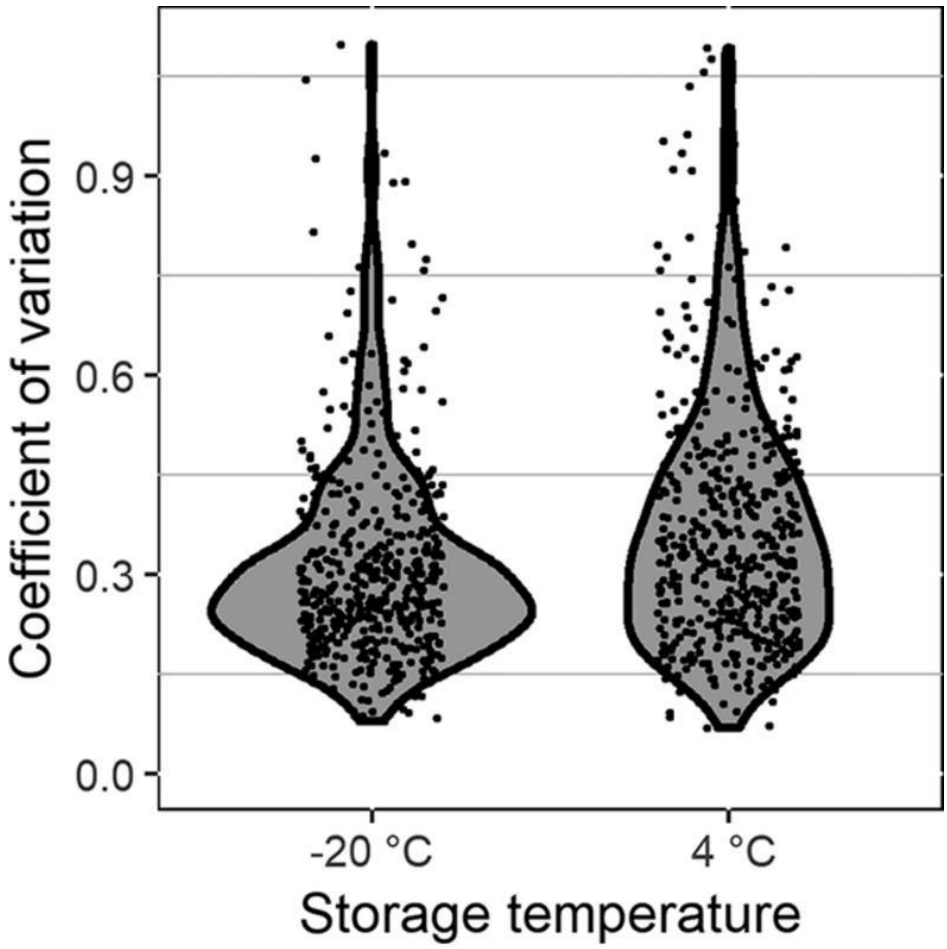


Figure 2.7. Violin plot of the coefficients of variation (COV: standard deviation / mean) of each metabolite when stored at 4°C or -20°C. The COV was calculated over time for each metabolite in each patient sample.

Table S2.1. PERMANOVA of Bray-Curtis distances between samples for each patient based on (A) storage temperature and (B) one versus two freeze-thaw cycles.

(A)

Patient 1	Df	Pseudo F	R2	P-value
Storage temperature	1	72.3	0.49	< 0.001
Residuals	74		0.51	

Patient 2	Df	Pseudo F	R2	P-value
Storage temperature	1	34.8	0.32	< 0.001
Residuals	74		0.68	

(B)

	Df	Pseudo F	R2	P-value
Storage temperature	1	0.37	0.005	0.72
Residuals	70		0.995	

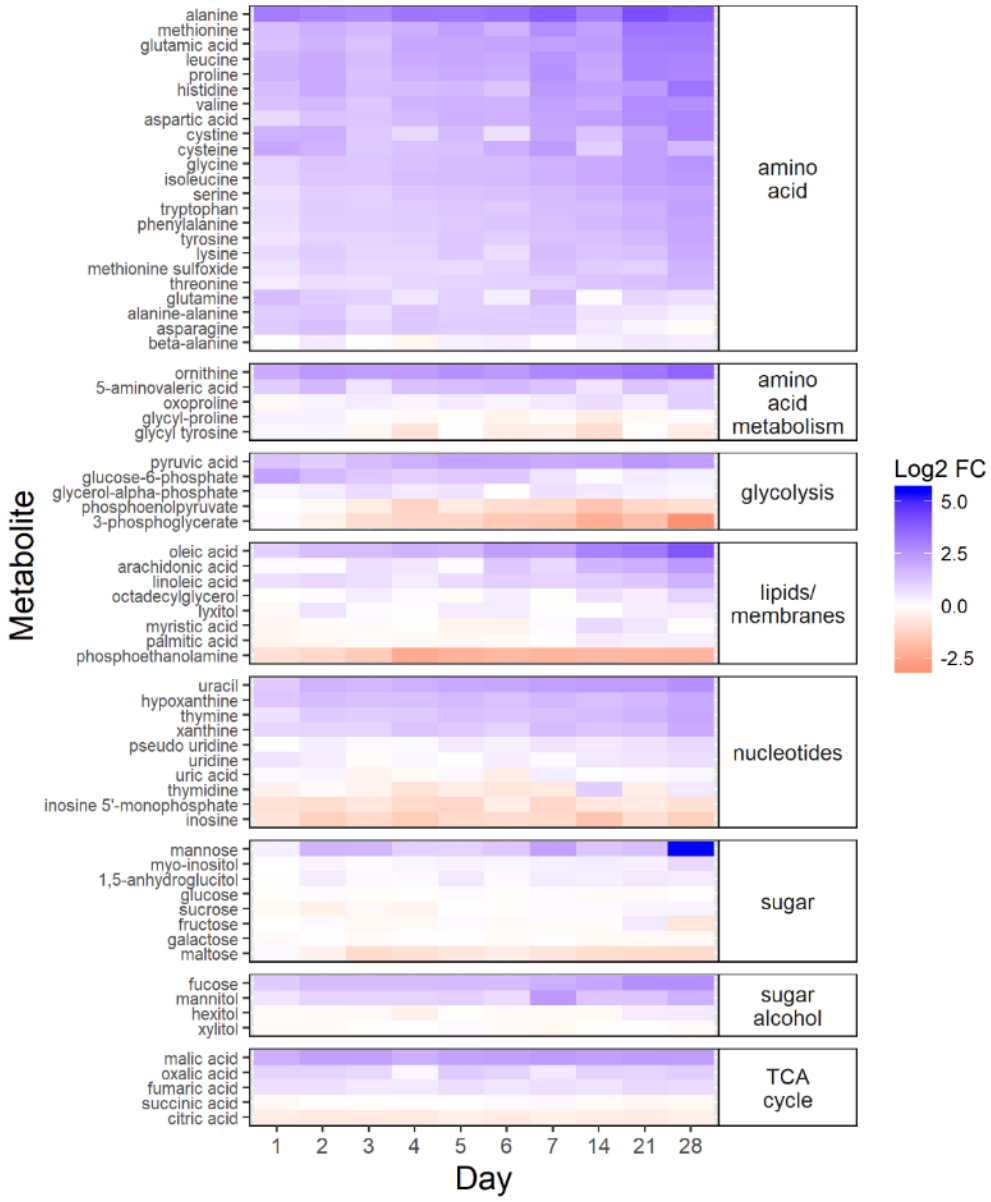


Figure S2.1. Heat map of a subset of metabolites detected by GC-MS after storage at 4 °C. Color scale represents the log base 2 of fold change in metabolite intensity at 4˚C compared to reference -80˚C.

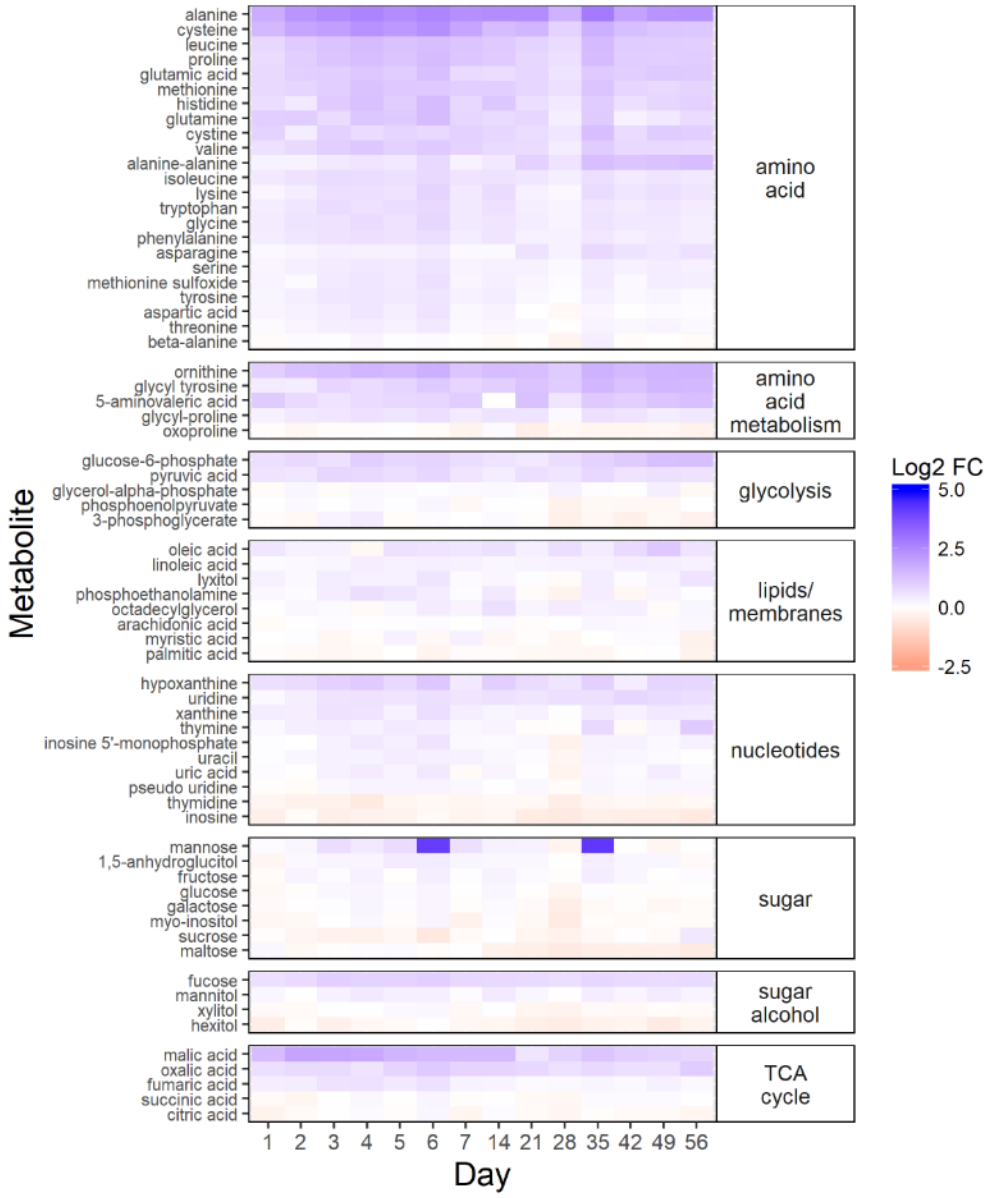


Figure S2.2. Heat map of a subset of metabolites detected by GC-MS after storage at -20 °C. Color scale represents the log base 2 of fold change in metabolite intensity at -20˚C compared to reference -80˚C.

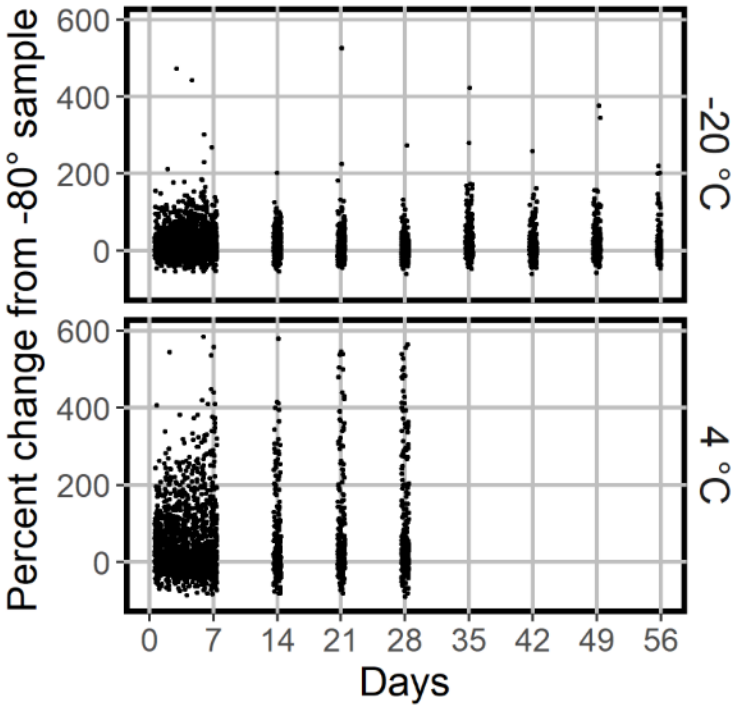


Figure S2.3. Average percent change of each of the 239 detected metabolites over time at 4 °C and -20 °C. Each point is the average percent change in both patient samples of a single metabolite.

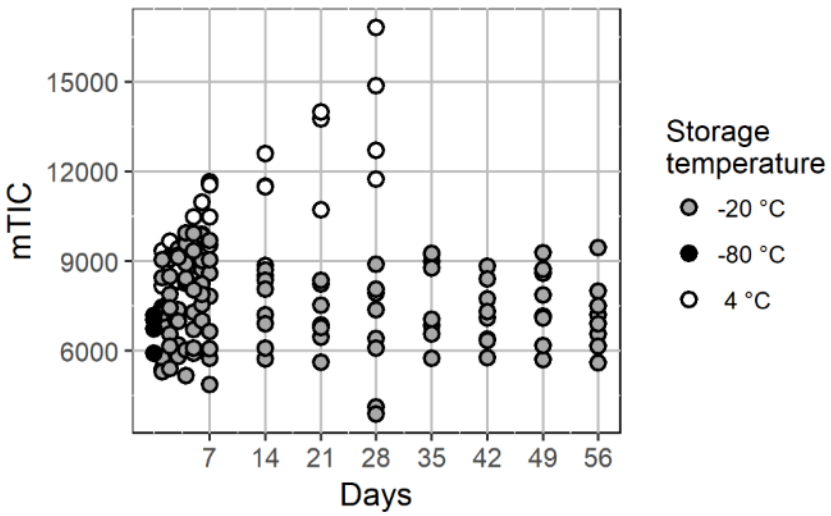


Figure S2.4. Total metabolite ion count (mTIC) of each sample over time colored by storage temperature.

CHAPTER 3

Predictable molecular adaptation of coevolving *Enterococcus faecium* and lytic phage EfV12-phi1

Co-authors: Andrew Oliver, Tara Gallagher, Claudia Weihe, Whitney England, Jennifer Martiny, and Katrine Whiteson

ABSTRACT

Bacteriophages are highly abundant in human microbiota where they coevolve with resident bacteria. Phage predation can drive the evolution of bacterial resistance, which can then drive reciprocal evolution in the phage to overcome that resistance. Such coevolutionary dynamics have not been extensively studied in human gut bacteria, and are of particular interest for both understanding and eventually manipulating the human gut microbiome. We performed experimental evolution of an *Enterococcus faecium* isolate from healthy human stool in the absence and presence of a single infecting Myoviridae bacteriophage, EfV12-phi1. Four replicates of *E. faecium* and phage were grown with twice daily serial transfers for eight days. Genome sequencing revealed that *E. faecium* evolved resistance to phage through mutations in the *yqwD2* gene involved in exopolysaccharide biogenesis and export, and the *rpoC* gene which encodes the RNA polymerase β' subunit. In response to bacterial resistance, phage EfV12-phi1 evolved varying numbers of 1.8 kb tandem duplications within a putative tail fiber gene. Host range assays indicated that coevolution of this phage-host pair resulted in arms race dynamics in which bacterial resistance and phage infectivity increased over time. Tracking mutations from population sequencing of experimental coevolution can quickly illuminate phage entry points along

with resistance strategies in both phage and host – critical information for using phage to manipulate microbial communities.

INTRODUCTION

Bacteriophages (phages) drive microbial diversity and function at both broad (Bouvier and Giorgio 2007) and fine scales (Mcshan et al. 2016) through their influences on bacterial community composition (Stern et al. 2012), and bacterial pathogenesis (Davis et al. 2000). Phages are estimated to be present at 10^9 virions per gram in the gut (Kim et al. 2011) and are therefore likely to have major influences on beneficial and pathogenic gut bacteria. Phages that lyse their host (lytic phages) or alter host virulence gene expression (some temperate phages) present a potentially rich pool of new therapies against antibiotic resistant pathogens (A. Wright et al. 2009; Oechslin et al. 2016). Recently, the clinical application of phages against highly antibiotic resistant bacteria (Chan et al. 2016; Schooley et al. 2017; Viertel, Ritter, and Horz 2014) has highlighted the need for well-controlled experiments that investigate the molecular interactions between phage and bacteria. Before phage-based therapies can be developed, we must have a solid understanding of how a targeted bacterial pathogen may evolve resistance to a treatment phage, and how the treatment phage responds to host resistance.

Reciprocal evolution of bacteria and phage, or coevolution (Thompson 1999), has been well-studied (Britt Koskella and Brockhurst 2014; Scanlan 2017; Martiny et al. 2014) in two model systems: *Pseudomonas fluorescens* and *Escherichia coli* (Britt Koskella and Brockhurst 2014; Bohannan and Lenski 2000; Hall et al. 2011). Although we can learn broad principles from these model systems, their study cannot replace experiments with more clinically relevant organisms to understand human associated phage-bacterial

interactions. We aimed to investigate coevolution in *Enterococcus faecium*, a common, but low-abundance member of the human gut microbiome that is also an important opportunistic pathogen. The World Health Organization classifies vancomycin-resistant *E. faecium* as a Priority 2 level pathogen in need of new antibiotic therapies (Lawe-Davies and Bennett 2017). Common enterococci infections include endocarditis, blood/wound infections, and urinary tract infections (Koch et al. 2004). Enterococci can also become dominating members of the gut community following antibiotic perturbation (Hendrickx et al. 2015), leading to dysbiosis and increased likelihood of infection (Van Tyne and Gilmore 2014). Developing a coevolution model using lytic phage and *Enterococcus* could therefore be a useful step towards addressing this global health threat. Coevolution experiments can quickly reveal candidates for the molecular basis of *Enterococcus*-phage interactions so that optimal cocktails of phages can be constructed. Indeed, cocktails of multiple phages with orthogonal infection mechanisms hold great promise as therapeutics (M. Yen, Cairns, and Camilli 2017a; Nale et al. 2018b).

The evolution of resistance to phage infection has been well documented and can happen through many routes. These include blocking phage adsorption through mutation, restriction-modification systems, CRISPR-Cas systems, and abortive infection (Dy et al. 2014). In addition, new mechanisms of phage resistance are still being discovered (Doron et al. 2018), which highlights the potential for discovery in the interactions between bacteria and phages. Coevolution between *Enterococcus* and its phages remains poorly studied, but resistance to one *Enterococcus* phage has been shown to evolve through mutation of an integral membrane protein to prevent phage adsorption (Duerkop et al.

2016). This remains one example, and *Enterococcus* may utilize an entirely different resistance mechanism during coevolution with a different phage.

We experimentally coevolved *E. faecium* with a lytic phage (EfV12-phi1) to characterize the genomic and phenotypic outcomes of their interaction. Phage EfV12-phi1 was isolated from sewage and has been previously referred to as "1" or "Φ1" (Jarvis, Collins, and Ackermann 1993). It is a member of the Twort-like family of Myoviridae phages, a group of strictly lytic phages that infect *Firmicutes* and generally demonstrate a broad host range. Closely related Twort-like phages have been previously employed for phage therapy and have demonstrated lethality against a long list of clinically relevant bacterial strains, including vancomycin-resistant enterococci (VRE); Group B, C, E, G *Streptococcus*; *Staphylococcus aureus*, and others (Klumpp et al. 2010; Khalifa et al. 2015b, 2018). The lysin of phage EfV12-phi1 has been previously shown to kill species of *Enterococcus* (including VRE), *Streptococcus*, and *Staphylococcus* (Yoong et al. 2004a).

We conducted four coevolution experiments where phage EfV12-phi1 was grown with *E. faecium* with 1:10 serial transfers twice daily, so that a large fraction of the population is carried over. To differentiate between genomic changes associated with coevolution versus those that might be due to laboratory adaptation, we compared these experiments to parallel control experiments where *E. faecium* was grown alone or phage EfV12-phi1 was propagated on a naïve host. Based on phage-host experiments in model systems, we expected to see mutations arise in the phage tail fibers that allow phage to recognize and bind their hosts and in bacterial surface receptors where phage often enter their hosts (Silva, Storms, and Sauvageau 2016).

RESULTS

E. faecium and phage EfV12-phi1 display arms-race coevolution dynamics

We coevolved *E. faecium* with lytic phage EfV12-phi1 as four replicate microcosms, passaging 16 times in eight days (every 12 hours), allowing for approximately 53 generations (**Figure 3.1**). Bacterial host control cultures were also set up in quadruplicate with identical conditions minus the phage. Quadruplicate phage controls were established by growing the phage on a naïve host, separating the phage from the host during each passage, and then adding the phage lysate to an independent aliquot of the naïve host control culture. Bacterial growth was monitored daily by optical density readings, which decrease when bacterial cells are lysed by phage. Phage infection initially reduced the density of all four bacterial cultures during the first day. This was followed by increased optical density after six to seven transfers (depending on the replicate), indicative of the evolution of resistance to phage (**Figure 3.2A**). In two replicate cultures, optical densities did not decline again after initial resistance arose, whereas in the other two replicate cultures, optical densities oscillated for the duration of the experiment.

At the final timepoint, bacterial populations in three of four replicates remained at a high optical density, despite relatively high concentrations of phage DNA (an approximation of phage abundance; **Figure 3.2B**). As expected, optical densities in the phage control cultures with naïve bacteria were consistently reduced upon infection by EfV12-phi1, and host control cultures (with no infecting phage) showed no reductions in optical density.

Ancestral and coevolved bacterial isolates were challenged with infection by ancestral and coevolved phages and bacterial lysis was scored using a plate-based assay (see Methods). These experiments showed that coevolved bacterial isolates (from the final

coevolution timepoint) were resistant to ancestral phage isolates, and the coevolved phage isolates infected ancestral bacterial isolates (**Figure 3.3**). In most cases, coevolved phages infected coevolved bacteria, suggesting that at least one round of coevolution had occurred (*E. faecium* evolved resistance, EfV12-phi1 in turn evolved an expanded host range to overcome this resistance). These results are consistent with arms race coevolutionary dynamics in which bacterial resistance and phage infectivity increase over time.

Resistance to phage evolves though exopolysaccharide and RNA polymerase mutations

To identify the bacterial mutations that led to resistance, and the phage mutations that enable infection of the freshly evolved host, we sequenced the populations from replicate microcosms at five timepoints (1, 4, 8, 12, and 16 transfers) from the coevolution treatment, three phage control timepoints (1,4,16) and two host control timepoints (1,16). These population reads were mapped to the ancestral *E. faecium* genome that was sequenced by both Illumina NextSeq and Oxford Nanopore MinION, yielding a high-quality reference genome in three contigs and one plasmid contig. Mutation frequencies for the population were calculated based on the percentage of reads supporting the mutant base divided by the total coverage. Non-synonymous mutations were not observed in host control bacteria but were observed in seven genes in coevolving populations. Many of these genes encode hydrolases and transferases (**Table 3.1**). Two genes were mutated in all four replicates: putative tyrosine kinase *yqwD2* and RNA polymerase B' subunit *rpoC*.

The putative tyrosine kinase, *yqwD2*, is involved in capsule exopolysaccharide production (**Figure 3.4A**). Replicates had different nonsynonymous mutations within this gene: three occurred on neighboring amino acid residues (P58H, P58L, G59V), while the fourth occurred twenty residues away (K89H) (**Figure 3.4A**). Mutations in the *yqwD2* gene

were first detected at timepoint eight and became more frequent in the coevolving bacterial populations over time. The increasing frequencies of different mutations in the same gene suggest convergent evolution toward a single mechanism for resisting phage infection.

The second bacterial gene observed to mutate when coevolving with phage EfV12-phi1 was the *rpoC* gene encoding RNA polymerase β' subunit (**Figure 3.4B**). A total of five different non-synonymous mutations were observed at high frequency in one or more replicates. The positions of mutations were mapped to the 3D structure of *E. coli's* RpoC, showing that all five mutations are located near each other on the interior portion of the protein near the active site (**Figure S3.1**).

Phage EfV12-phi1 combats resistance through tandem tail fiber duplications

Mutations in the phage genome were also tracked over time as the phage coevolved with the host bacteria. Four phage genes mutated throughout the experiment. Three of these mutations also occurred in all replicates of the phage controls, indicating that they are likely to generally increase infectivity for this specific host and are not a response to the evolution of bacterial resistance. One of the phage-control mutations occurred in a putative structural capsid gene and resulted in a change from asparagine to lysine. In all replicates, this mutation started at a low frequency at transfer 4 (the first sequenced time point) and increased in frequency over time (**Table 3.2**). The other two genes encoded hypothetical proteins that were deleted from the genome between timepoints 8 and 12 (**Table 3.2**). These genes are located next to each other and are near the several terminally redundant repeats EfV12-phi1 uses to circularize its genome, suggesting a likely mechanism for excision of these genes.

The coevolution-specific phage mutation occurred in a gene encoding a putative tail fiber. Partial duplications of this gene occurred in all four coevolution replicates and never in the phage controls. Specifically, a 1.8 kb segment of a 6.6 kb putative tail fiber gene underwent in-frame tandem duplications (**Figure 3.5**). Over time, replicates acquired varying numbers of duplications 400 bp upstream of a predicted carbohydrate-binding domain. The duplication was initially observed as an increase in sequencing coverage present in all four populations beginning between transfers four and eight and persisting until the end of the experiment (**Figure 3.5a,c**). PCR was performed with primers were flanking the entire duplication so that the amplicon would increase in size if duplications occurred. For coevolved phage populations (**Figure 3.5b**) and isolates (data not shown), multiple amplicons of increasing size were observed that represent the size of the original tail fiber gene as well as larger tail fiber genes that contain duplications.

MinION long-read sequencing was performed on a phage isolate from the final timepoint of population 4 to resolve the duplication. Of the 1,021 reads spanning the entire tail fiber gene, 134 reads had no duplication (the original tail fiber gene), 852 reads had one duplication (two tandem copies of the duplicated sequence), 32 reads had two duplications, two reads had three duplications, and one read had four duplications. Thirty-four reads were found to consist of only tandem copies of the duplicated 1.8 kb sequence, ranging from 4 to 11 copies (7 kb to 20 kb in length). The mechanism by which these tandem duplications altered phage infectivity is currently not known; the duplication did not appear to be a diversity-generating mechanism, as only a single replicate acquired a SNP within the duplicated region (Table 2). The duplications were first detected at transfer eight, after a dramatic increase in bacterial abundance, which we attribute to the evolution

of resistance to phage infection. The timing and exclusive occurrence in the coevolution treatment suggests that these tail fiber duplications were a phage response to the bacterial evolution of resistance.

DISCUSSION

To our knowledge, this represents the first effort to characterize phage-bacteria coevolution in *Enterococcus* - a common commensal in the gut microbiome that is also an important opportunistic pathogen. Similar to other well characterized systems, the experiments revealed coevolutionary arms race dynamics between *E. faecium* and its phage involving mutations in phage tail fibers and bacterial surface structures. They further demonstrated parallel coevolution among replicates and therefore predictable molecular adaptation. In particular, we identified bacterial exopolysaccharide mutations suggestive of hindering phage adsorption and RNA polymerase β' subunit mutations with the potential to disrupt the phage replication cycle. However, we also identified what appears to be an unknown phage escape strategy involving large tandem repeats in the tail fiber gene. While some of the basic dynamics and molecular mechanisms of coevolution appear to be similar across many phage-host pairs (Labrie, Samson, and Moineau 2010; Samson et al. 2013a), experimental coevolution in this understudied system allowed us to quickly identify unique adaptation strategies.

Coevolving bacteria acquired mutations in the *yqwD2* gene, and we hypothesize that these mutations are at least partially responsible for the resistance phenotype seen in *E. faecium* coevolving with phage EfV12-phi1. The *yqwD2* gene is part of a capsule production operon that is well conserved among Firmicutes; it is known as Ywq in *Bacillus subtilis*, Cps in *Streptococcus pneumoniae*, and Eps in *Streptococcus thermophilus* (Stingele et al. 1996;

Bentley et al. 2006; K. L. Palmer et al. 2012). In *Streptococcus thermophiles*, the *epsD* gene (35% amino acid identity to *E. faecium yqwD2*) encodes a cytoplasmic tyrosine kinase that regulates the activity of EpsE, a phosphogalactosyltransferase. Disruption of either *epsD* or *epsE* abolished extracellular polysaccharide synthesis (Minic et al. 2007). Mutations in exopolysaccharide production genes have been shown to inhibit phage infection in *Enterococcus faecalis* (F. Teng et al. 2009), *Lactococcus lactis*(Forde and Fitzgerald 1999). Interestingly, two of these mutations occurred at residue 58 and one at residue 59 which are the beginning of a conserved nucleotide binding motif (GEGKS) (Stingele et al. 1996). A homologous protein structure within the conserved domain database (CDD) shows that this region of the protein is highly accessible. In line with protein models previously proposed (Stingele et al. 1996), perhaps these mutations interfere with the function of YqwD2, subsequently altering the structure, length, or quantity of exported exopolysaccharides (Morona, Van den Bosch, and Manning 1995; Bastin et al. 1993; Minic et al. 2007). While the phage receptor of phage EfV12-phi1 is unknown, distantly related phages *Staphylococcus* phage K and Bacillus phage SP01 bind to cell wall teichoic acids (Estrella et al. 2016; Yasbin, Maino, and Young 1976). Similarly, phage EfV12-phi1 may bind to certain motifs of exopolysaccharides, so that modification of exopolysaccharides hinders phage adsorption. Further, several bacterial mutations that were not conserved among all replicates encoded sugar metabolism and modification functions which could also alter the structure and modifications present on exopolysaccharides. Future genetic knockout experiments will be useful in determining the degree to which these mutations confer resistance.

Coevolving bacteria also acquired mutations in the *rpoC* gene, which encodes the RNA polymerase β' subunit. Phage EfV12-phi1 does not encode its own RNA polymerase, so it needs to interact with the host RNA polymerase both to shut down transcription of host genes and to transcribe phage genes. Mutations in the RNA polymerase *rpoC* gene could be a mechanism to resist phage infection by disrupting RNA polymerase activity. Phages produce proteins to bind or modify host RNA polymerase subunits, including the β' subunit, to shut down host transcription and increase affinity for phage DNA (Mailhammer et al. 1975; Nechaev and Severinov 2003; Hodgson, Shapiro, and Amemiya 1985; Hesselbach and Nakada 1977). The mutation of residues that are modified or bound by phage proteins during infection could be a mechanism by which *E. faecium* can resist infection by phage EfV12-phi1. Five of the six different *rpoC* mutations observed were unique to single replicates, but all are located near each other in the 3D structure of RpoC, which suggests they all provide resistance to phage EfV12-phi1 through a common mechanism. The mutations in RpoC are localized similarly as the mutations that arise with the genetic suppressors of a protein, DksA, that regulates *Escherichia coli* RpoC in response to nutrient availability (Rutherford et al. 2009). This suggests that the bacterial resistance arises through a general RpoC suppression mechanism that reduces phage success although it may not be driven by direct interaction with between phage proteins and RpoC.

The only phage mutations unique to coevolution (and not present in the evolution of phage EfV12-phi1 to naive host) were tandem duplications within a putative tail fiber gene. Myoviruses have short and long tail fibers, the latter of which are responsible for scanning the host cell surface and identifying the receptor. This gene has been confirmed to be the long tail fiber in a closely related phage, phiEF24C (Uchiyama et al. 2011). A point mutation

in the homologous tail fiber of phiEF24C was seen to increase adsorption to several strains of *E. faecalis*. The duplication observed in this experiment occurs in a region of the gene that differs between EfV12-phi1 and phiEF24C. Protein homology analysis of the gene indicates a predicted carbohydrate-binding domain 100 nucleotides downstream from the duplicated region, but no conserved domains were predicted within the duplication itself. The duplicated region does not appear to generate sequence diversity which might allow recognition to different bacterial surface receptors, as has been observed in phage λ (Meyer et al. 2012). The timing of EfV12-phi1 tandem duplications suggests that they are a response to the evolution of bacterial resistance to phage infection. Overall, we speculate that phage may respond to bacterial capsule changes through modifications in the tail fiber. Although mutations in the tail fibers are common mechanisms by which phages adapt to modified bacterial receptors (Tétart et al. 1996; Scanlan et al. 2011), examples of duplications as large as the one seen in this study (1.8 kb per duplication) have not been seen before.

Phage therapy has long been a proposed solution to the growing problem of antibiotic resistant bacteria, with recent successful cases of phage therapy in the US following a compassionate use exemption. However, phage therapy is limited by a lack of well characterized phages infecting human pathogens (Schooley et al. 2017; Zhvania et al. 2017; Duplessis et al. 2017). Phage therapy utilizes phage cocktails, which include a mix of different phages with orthogonal targets to counter the evolution of bacterial resistance. Understanding the dynamics and outcomes of bacteria-phage interactions using experimental coevolution would facilitate phage cocktail design. For example, EfV12-phi1 has broad host-range and selects for *E. faecium* mutations related to exopolysaccharide

synthesis, suggesting that a cocktail including EfV12-phi1 would be most effective if the other cocktail phages targeted host structures other than exopolysaccharide.

Phage EfV12-phi1 may have therapeutic potential, given that it is widespread and the host range was previously (Yoong et al. 2004a) found to include a wide range of pathogens. Predictability of phage-host interactions is desirable to ensure safety of phages and for phage cocktail design. In this phage-host pair, we observed consistent outcomes from all four replicates, despite the stochasticity of mutations that lead to those outcomes. Nine of the eleven observed bacterial mutations were not shared among all replicates, but the functions encoded by these genes shared similar features (hydrolases, transferases, sugar metabolism/modification).

In these experiments, in just 8 days, we quickly identified phage and host genes that are under selection during coevolution. Experimental manipulation of phage-host interactions, and periodic tracking of their mutational trajectories, offers exceptional insight into the mutational arms race – beyond traditional sequencing and annotation efforts. While coevolution in artificial laboratory conditions may not be reflective of coevolution that happens in a natural environment, learning about the potential outcomes of coevolution provide useful information. As microbial culturing and enumeration becomes increasingly automated, a large number of phage-host interactions can be tested in order to thoroughly investigate the mechanism of phage-host co-evolution in a diversity of clinically relevant hosts. Such insights are critical to the eventual development of phage therapies for clinical use.

MATERIALS AND METHODS

Bacterial strains and phage

The bacteria used in this study was *Enterococcus faecium*, Strain TX1330, HM-204, was isolated from healthy human feces and obtained through BEI Resources, NIAID, NIH as part of the Human Microbiome Project. The phage used for this study was *Enterococcus* Phage EfV12-phi1, isolated on *E. faecalis,* from Canadian sewage in 1975 (HER number 339; d'Herelle collection, Laval University, Quebec, Canada).

Coevolution of *Enterococcus* bacteria and phage

A culture of *E. faecium* TX1330 growing exponentially (OD600 = 0.3) in brain heart infusion (BHI) broth was split into twelve replicates of 10 mL culture in 15 mL Falcon tubes. Four replicates were designated bacterial host control, four were phage control, and four were coevolution. Phage EfV12-phi1 was added to the coevolution and phage control cultures at an MOI of approximately 0.003. Cultures were incubated shaking with loose caps at 37°C. Every 12 hours for eight days, 1 mL of the replicate host control and coevolution tubes were inoculated into 9 mL of new BHI broth. For the phage control, 1 mL of phages was separated from bacteria by syringe filtration through a 0.2 um polyethersulfone filter (GE Healthcare Life Sciences) and mixed with 1 mL of the contemporary host control in 8 mL of new BHI broth. Phages were separated by centrifugation at 12,000g and filtering the supernatant using 0.2 um filter. Performing 1:10 dilutions at each passage, we estimate 3.5 generations (doublings) are required to reach stationary phase again, resulting in approximately 56 generations total. After each dilution, 900 uL of 12-hour culture containing the population of bacteria and phages was added to 600 uL of 50% glycerol and stored at -80°C. Transfer numbers 1,4,8,12, and 16 were chosen for sequencing. At each timepoint (after 12 hours of growth), the OD600 of each culture was measured.

Host range assay

Host range of phage and bacterial isolates from *E. faecium* were determined with streak assays as described before (Harcombe and Bull 2005). Bacteria and phage were isolated from the first and last timepoints of the host control and coevolution replicate populations one and two. Bacteria were isolated by streak plating and picking single colonies. Phages were isolated by performing a double agar overlay with 100 ul of the raw population and then picking plaques (so that phages are growing on contemporary hosts from the same replicate population). Phages were amplified by performing plaque assays on ancestral hosts and harvested by soaking plates in 5 mL SM buffer followed by filtration of collected SM buffer through a 0.2 um syringe filter. 20 ul of each bacterial isolate and 20 ul of each phage isolate was streaked perpendicularly across an agar plate. The intersection of the bacteria and phage was examined and scored for lysis. In total, three ancestral hosts and 12 coevolved hosts were crossed against six ancestral phages and 16 coevolved phages.

DNA extraction, library preparation, and sequencing

DNA was extracted from the populations of bacteria and phages in the chosen timepoints with the Zymo Universal DNA extraction kit using the recommended protocol provided by the manufacturer. Sequencing libraries were prepared with Illumina's Nextera kit using methods outlined in Baym et al. (Baym et al. 2015). The libraries were loaded onto an Illumina Next-Seq at 1.8 picomolar concentration using Illumina's mid-output kit for 75 bp paired end sequencing.

A more complete bacterial reference genome was assembled using Oxford Nanopore's 1D Genomic DNA Ligation kit (Goodwin, Wappel, and McCombie 2017). Briefly,

DNA was repaired using the FFPE DNA repair kit (New England Biolabs) and cleaned up using AMPure XP beads (Beckman Coulter). The repaired DNA was dA-tailed using NEBNext Ultra End Repair (New England Biolabs) and sequence adapters were ligated using Blunt TA ligase master mix (New England Biolabs). The MinION sequencer was primed, per manufacturer's instructions, and 700ng of DNA was loaded onto the sequencer. The run was allowed to generate data for 48 hours. Sequence data from the MinION and Illumina sequence data from timepoint one of the host control were used together to generate a host reference genome using the MIRA assembler (Chevreux, Wetter, and Suhai 1999).

Genome assembly and annotation

The reference genome for *E. faecium* TX1330 was assembled using reads from time point 1 of the host control. Reads were assembled using the PATRIC smart assembler (Wattam et al. 2017), which combines the two best assemblies from SPAdes (Bankevich et al. 2012), IDBA (Peng et al. 2010), and Velvet (Zerbino and Birney 2008) assemblers. The phage was assembled using SPAdes (Bankevich et al. 2012). The resulting contigs were annotated using PATRIC's annotation pipeline, which uses RASTtk for gene calls (Wattam et al. 2017). The sequenced genome of *E. faecium* TX1330 can be found at GenBank: QYBD00000000.1, and the EfV12-phi1 genome can be found at GenBank: MH880817.

Genomic mutation analysis

Paired-end reads were run through Breseq (Deatherage and Barrick 2014) once using the ancestral phage EfV12-phi1 as the reference genome and once using the ancestral *E. faecium* TX1330 with default parameters. Briefly, Breseq uses Bowtie2 (Langmead and Salzberg 2012) to align reads to a reference genome and creates a SAM file which

SAMTOOLS converts to a pileup file. Custom R scripts were then used to parse through the resulting alignments and detect mutations at greater than 10% frequency. Mutations were labeled as synonymous or non-synonymous by Breseq, and all predicted non-synonymous mutations were manually investigated using Geneious (Biomatters v9.0). All bacterial mutations were visualized using Geneious and phage mutations using Geneious and ggPlot2 in R.

Tail fiber PCR

Phage populations from the final timepoint of all four replicates were grown by adding 10 ul of the frozen timepoint 16 cultures to 10 mL BHI and grown overnight shaking at 37 °C. Cultures were then spun down and the supernatant was syringe filtered through a 0.2 um polyethersulfone filter. Phages were then concentrated down to 1 mL using Amicon 100 kDa centrifugal filter units. DNA was extracted from concentrated phages using Zymo Universal DNA extraction kit.

Primers were designed outside the duplicated region of the tail fiber gene so that the amplicon would be longer if the region was duplicated. The primers used were F: 5' TGTTGCACCAGAAAACGCAG 3' and R: 5' AGGTCTGTACGAGCCGTGTA 3'. PCR was run using Phusion polymerase with the following protocol: 98 °C: 30 seconds, (98 °C 10 seconds, 53 °C 30 seconds, 72 °C 10 minutes) x 35, 72 °C 10 minutes. Amplicons were visualized on a 1 % agarose gel using Invitrogen SYBR gel stain.

Location of mutations in RNA polymerase B' structure

The structure of the *E. coli* RNA polymerase B' subunit was downloaded from Protein Data Bank (4JK1, DOI: 10.2210/pdb4JK1/pdb). The amino acid sequence of the *E. faecium* TX1330 RNA polymerase B' subunit was aligned to the *E. coli* sequence to find the

corresponding locations. The structure and locations of mutations were visualized using PyMOL (Schrödinger, LLC 2015).

MinION sequencing of tail fiber duplication

The Oxford Nanopore MinION sequencer was used to sequence a phage isolate that contained the tail fiber duplication. The phage was isolated by picking a plaque from timepoint 16 of replicate 4 (directly plating 100 ul of the population). The phage isolate was propagated on a contemporary (final timepoint) *Enterococcus* isolate from population 4 to get enough phage DNA for sequencing. DNA was extracted using a Zymo Quick-DNA micro kit. DNA was prepared for MinION sequencing according to manufacturer's recommendations using the 1D Genomic DNA by ligation protocol as described above. A total of 199,734 reads were generated with a median sequence length of 3,057 bp. Bowtie2 was used to extract the 57% of reads that aligned to the phage genome; the remaining reads aligning to the bacterial genome were discarded. The data was analyzed in Geneious to determine the number of duplications in the tail fiber gene. A total of 5,400 reads aligned to the tail fiber gene and were over 3 kb so could span the length of a single duplication (1.8 kb).

ACKNOWLEDGEMENTS

We would like to acknowledge Dr. Heather Maughn for thoughtful comments and edits, Dr. Ali Mortazavi and his lab for allowing us to use their NextSeq, and Dr. Rich Puxty for guidance on analyzing our sequencing data. Thanks to all the members of the Whiteson Lab for providing advice and feedback throughout this project.

DATA AVAILABILITY

All sequencing data has been deposited to the SRA at PRJNA490385 The genome for phage EfV12-phi1 can be found at GenBank: MH880817 and our assembly of *Enterococcus faecium* TX1330 can be found at Genbank assembly accession: GCA_003583905.1.

FIGURES AND TABLES

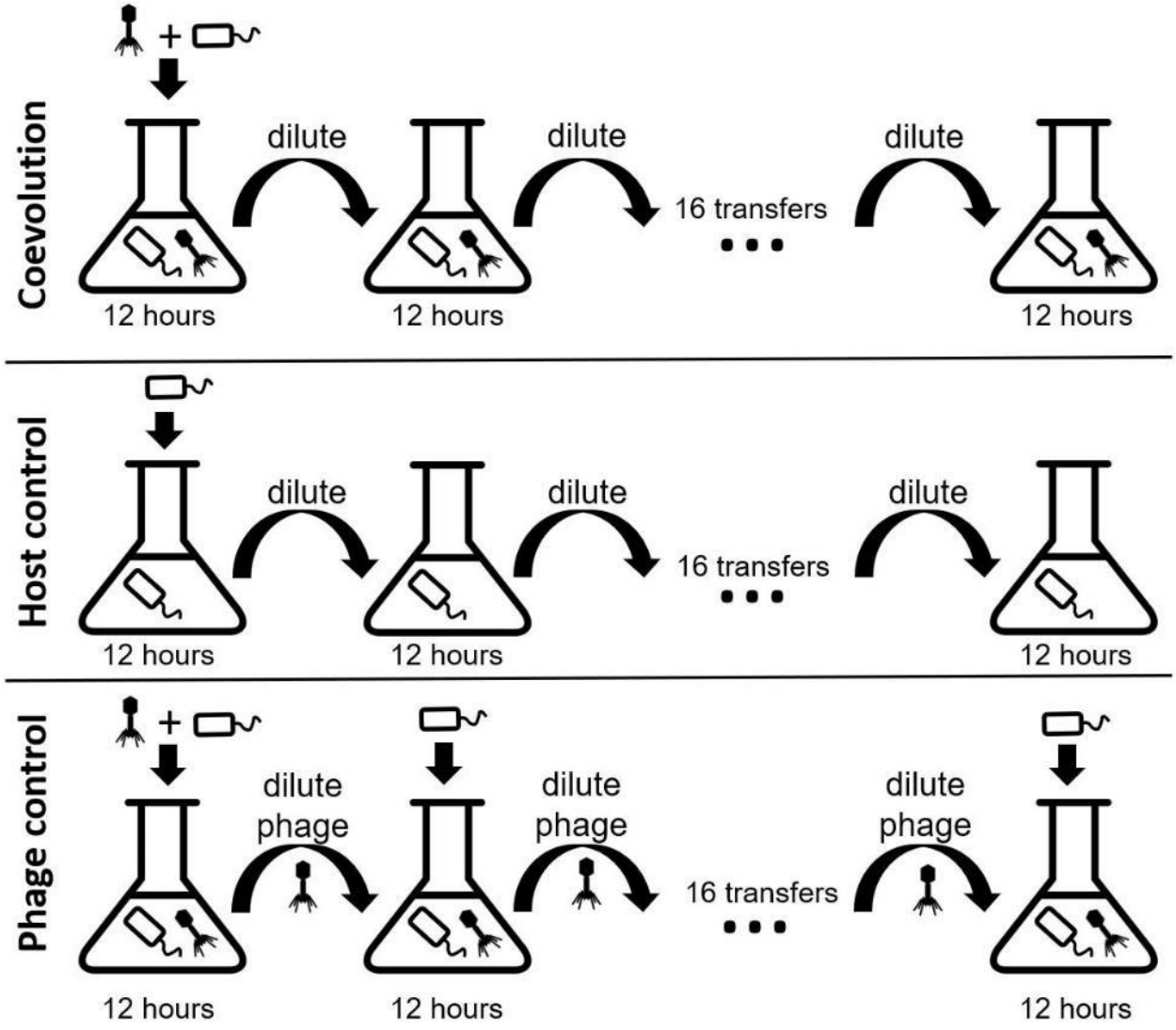


Figure 3.1. Experimental design of the three branches of the study. In each branch, phage and bacteria or only bacteria were added to a microcosm and allowed to grow for 12 hours before being diluted 10-fold. The phage control filtered out the bacteria during each dilution, preventing bacterial coevolution. Each branch was done in quadruplicate.

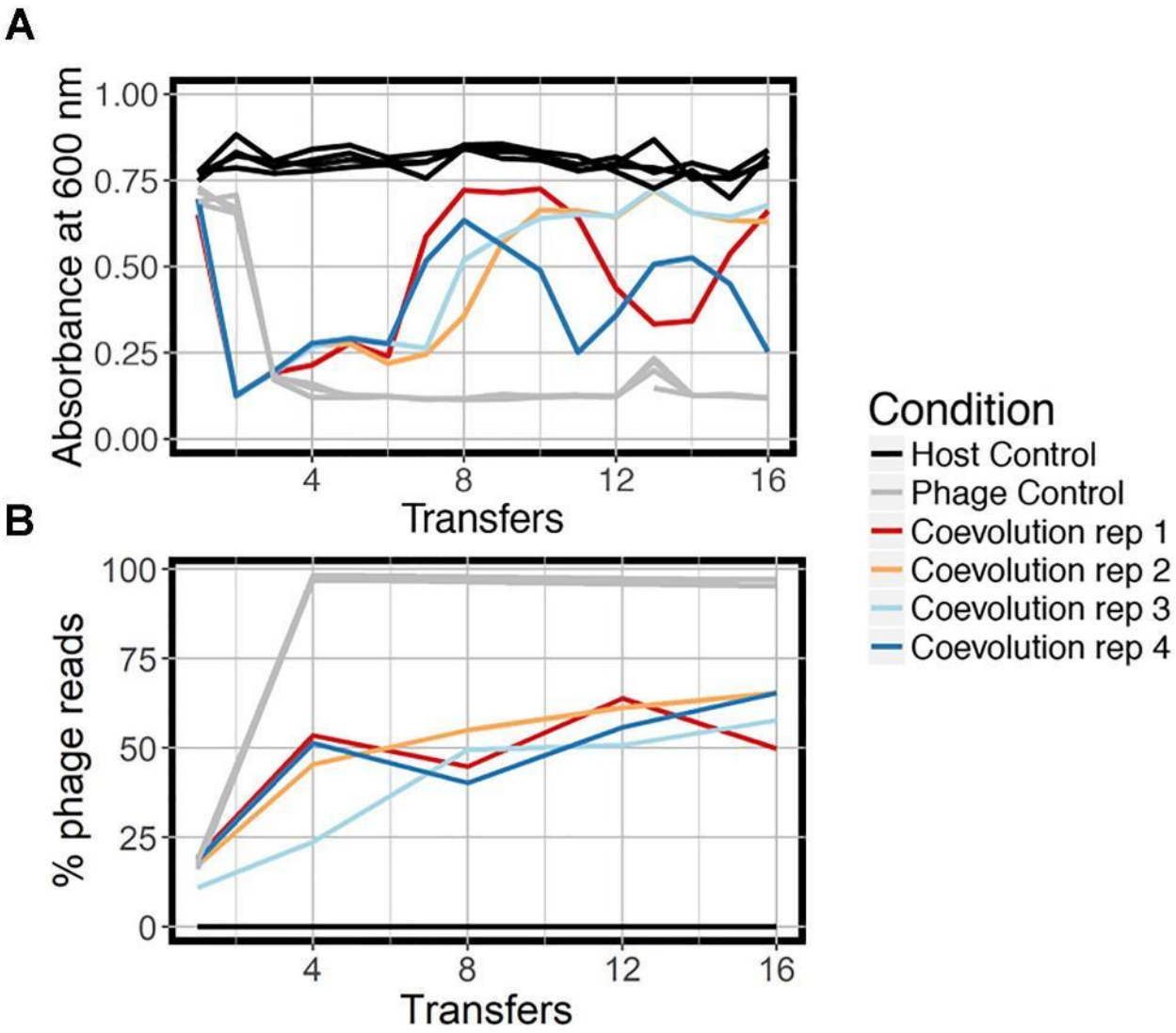


Figure 3.2. Growth dynamics of experimental coevolution. A) Optical density of the bacteria for each branch of the experiment, measured at the end of 12 hours prior to diluting back 10-fold in fresh BHI media. All replicates of the host control and phage control are shown in the same color because there was little variation. B) Proportion of total sequenced reads mapping to phage EfV12-phi1 indicates the relative abundance of this phage at each timepoint. Reads that did not map to phage mapped to *E. faecium*.

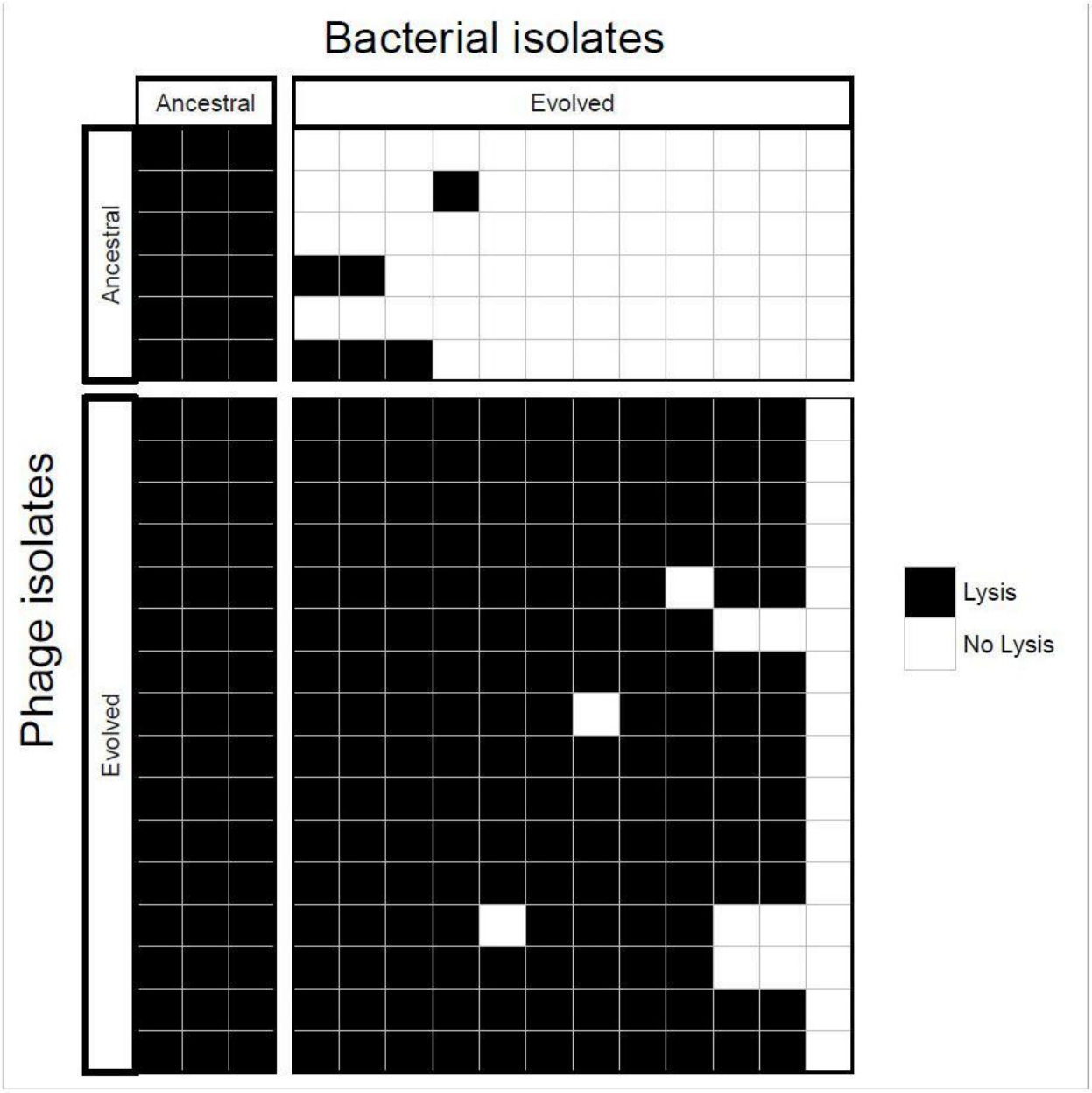


Figure 3.3. Host range analysis of phage and bacterial isolates. Bacteria and phage isolated from the initial and final timepoints were tested for infectivity. Each box represents whether lysis occurred when a single phage isolate crossed with a single bacterial isolate.

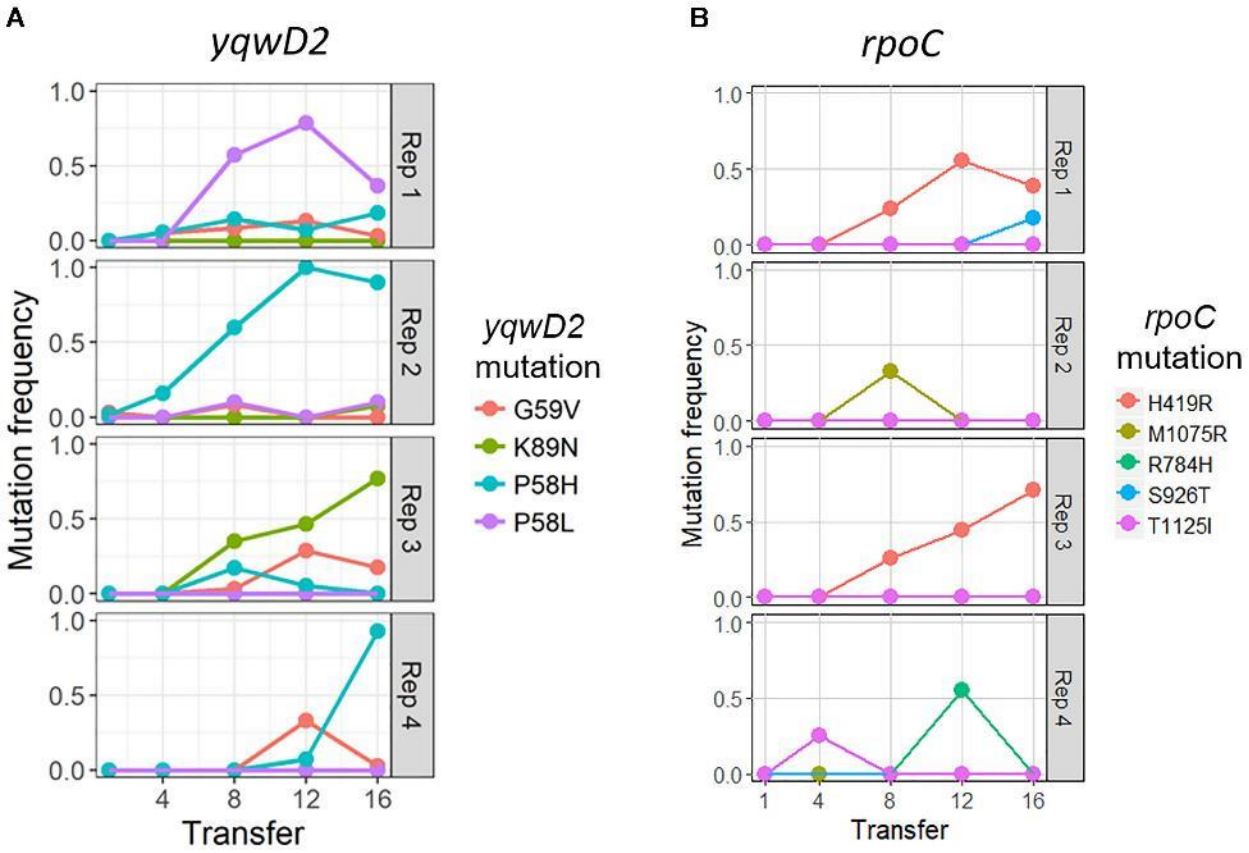


Figure 3.4. Frequency of common bacterial mutations over time. Population frequency of mutations in **A)** capsule biosynthesis tyrosine protein kinase *yqwD2* and **B)** RNA polymerase B' subunit gene *rpoC*. All mutations present at a frequency of 10% in one timepoint in one replicate shown.

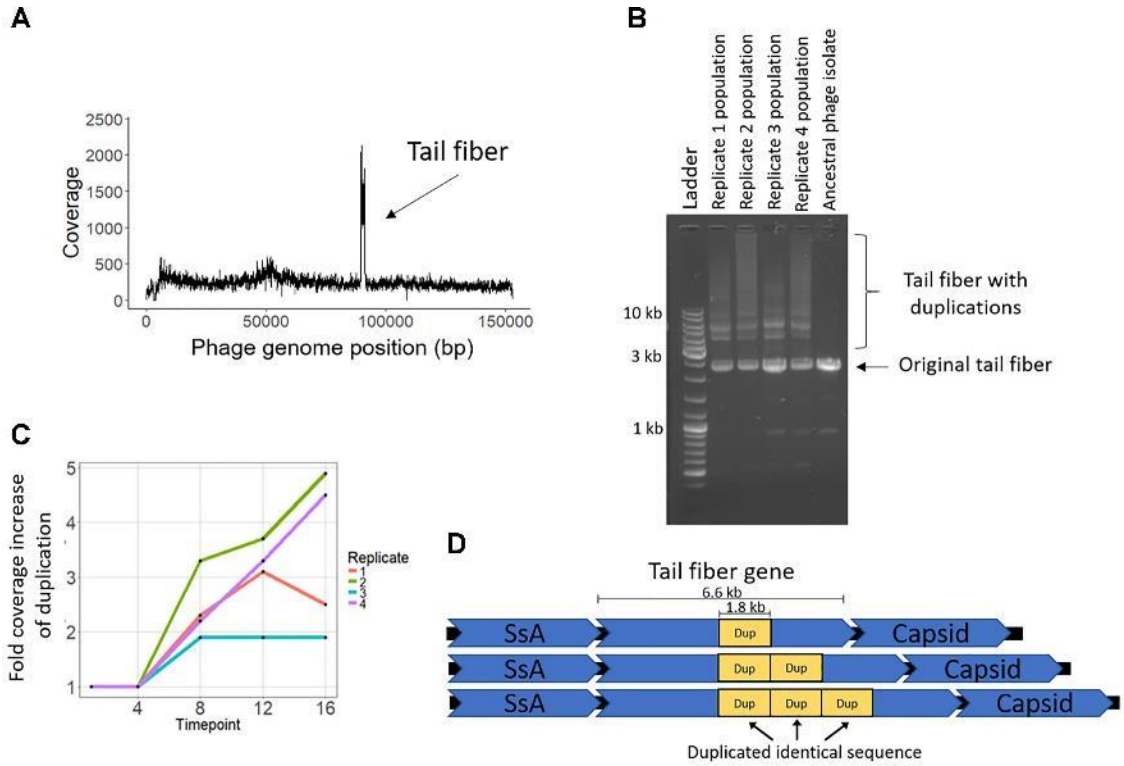


Figure 3.5. Phage EfV12-phi1 evolved tandem duplications in the tail fiber gene to increase its infectivity. **A)** Average coverage along the phage genome for the phage population of replicate 2 at the final timepoint. Duplication was first noticed by this spike in sequencing coverage. Reads were mapped to the original phage genome so the the duplication in the tail fiber appears as a spike in coverage. **B)** Duplications in the tail fiber visualized by PCR using primers that flank the tail fiber. Duplications resulted in a larger amplicon. Each replicate population at the final timepoint is shown as well as the ancestral phage. **C)** Presence of tail fiber duplications over time shown by the fold coverage increase in the duplicated region divided by the average coverage of the rest of the phage genome. **D)** Schematic of the phage tail fiber tandem duplication within the gene. Reads spanning the tail fiber gene containing up to three duplications (four total copies of duplicated sequence) were seen with MinION long read sequencing.

Table 3.1. All mutations present in *E. faecium* TX1330 at the final timepoint.

All bacterial mutations are from timepoint 16 of coevolution replicates . No mutations were seen in the host controls. Locus tags for each gene can be found in **Table S1**.

Replicate	Gene / Predicted function	Type	AA change	Frequency (%)
1	*RpoC*	nonsynonymous snp	H419R	38.2
1	*RpoC*	nonsynonymous snp	S926T	21.1
1	*RpoC*	nonsynonymous snp	L800V	14.8
1	Hypothetical protein in capsule synthesis locus	nonsynonymous snp	M29I	30
1	*yqwD2*	nonsynonymous snp	P58L	37.6
1	*yqwD2*	nonsynonymous snp	P58H	19.2
2	Malonate decarboxylase beta subunit / Malonate decarboxylase gamma subunit CDS	nonsynonymous snp	G148V	45.5
2	Predicted hydrolase of the HAD Superfamily CDS	nonsense	S191stop	33.3
2	murA - UDP-N-acetylglucosamine 1-carboxyvinyltransferase	nonsynonymous snp	G20C	40
2	*yqwD2*	nonsynonymous snp	P58H	100
3	*RpoC*	nonsynonymous snp	H419R	79.2
3	*yqwD2*	nonsynonymous snp	K89N	72.2
3	hydrolase, haloacid dehalogenase-like family CDS	nonsense	E68stop	50
4	*yqwD2*	nonsynonymous snp	P58H	92.5

Table 3.2. All mutations present in phage EfV12-phi1 at the final timepoint.

All phage mutations are from timepoint 16 in the coevolution and phage control populations. Frequencies given as population frequency for snps and fold coverage increases/decreases in the population for duplications and deletions. Locus tags correspond to Genbank record MH880817.

Replicate	Condition	Gene / Predicted function	Locus tag	Type	AA change	Frequency
1	Coevolution	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 375x coverage
1	Coevolution	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 545x coverage
1	Coevolution	Tail fiber	EFV12PHI1_98	tandem duplication	-	+ 3x coverage
1	Coevolution	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	99 %
2	Coevolution	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 250x coverage
2	Coevolution	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 58x coverage
2	Coevolution	Tail fiber	EFV12PHI1_98	tandem duplication	-	+ 5x coverage
2	Coevolution	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	99 %
3	Coevolution	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 33x coverage
3	Coevolution	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 30x coverage
3	Coevolution	Tail fiber	EFV12PHI1_98	tandem duplication	-	+ 3x coverage
3	Coevolution	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	77.5 %
4	Coevolution	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 896x coverage
4	Coevolution	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 896x coverage
4	Coevolution	Tail fiber	EFV12PHI1_98	tandem duplication	-	+ 5x coverage
4	Coevolution	Tail fiber	EFV12PHI1_98	non-synonymous snp	R1460H	23.9 %
4	Coevolution	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	93.5 %
1	Phage control	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	81.3 %
1	Phage control	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 5x coverage
1	Phage control	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 5x coverage

2	Phage control	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	90.3 %
2	Phage control	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 50x coverage
2	Phage control	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 48x coverage
3	Phage control	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	92.1 %
3	Phage control	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 14x coverage
3	Phage control	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 12x coverage
4	Phage control	Capsid and Scaffold	EFV12PHI1_97	non-synonymous snp	N306K	86.8 %
4	Phage control	Hypothetical Protein 8	EFV12PHI1_123	whole gene deletion	-	- 20x coverage
4	Phage control	Hypothetical Protein 9	EFV12PHI1_126	whole gene deletion	-	- 18x coverage

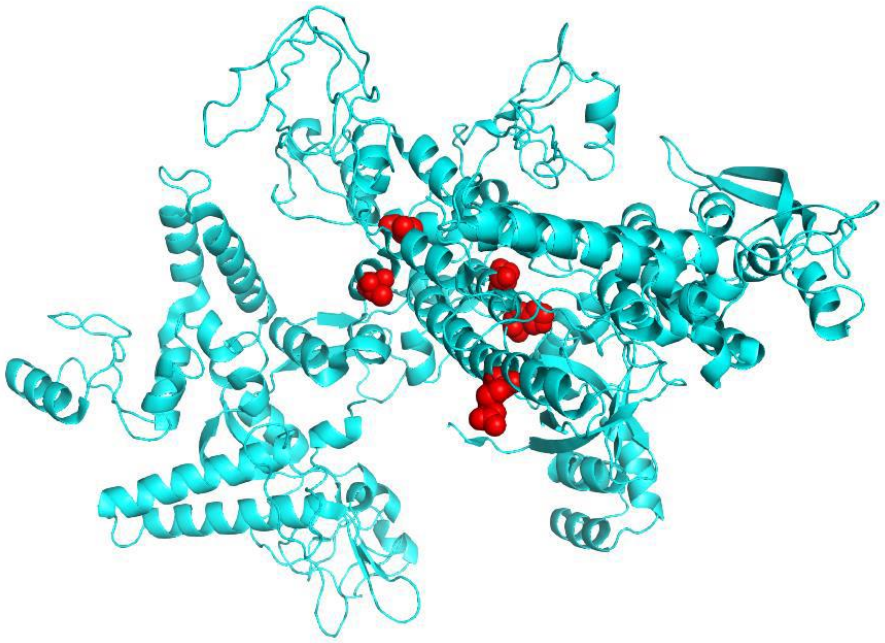


Figure S3.1. Structural representation of RpoC with the positions of the mutations highlighted in red. Structure of *E. coli* RNA polymerase B' shown with corresponding mutations in *E. faecium* TX1330 RNA polymerase B' highlighted.

Table S3.1. Locus tags for genes mutating in *Enterococcus faecium* TX1330.

Gene	**Locus tag GCA_000159675.1**	**Locus tag GCA_003583905.1**
RpoC	HMPREF0352_2730	D3Y30_03855
Histidinol-phosphatase (EC 3.1.3.15) CDS	HMPREF0352_0179	D3Y30_13030
yqwD2	HMPREF0352_1902	D3Y30_11120
Malonate decarboxylase beta subunit / Malonate decarboxylase gamma subunit CDS	HMPREF0352_0928	D3Y30_01740
Predicted hydrolase of the HAD Superfamily CDS	HMPREF0352_0295	D3Y30_06415
murA - UDP-N-acetylglucosamine 1-carboxyvinyltransferase	HMPREF0352_0323	D3Y30_06555
hydrolase, haloacid dehalogenase-like family CDS	HMPREF0352_1574	D3Y30_13500

CHAPTER 4

Broad host range *Brockvirinae* phages infect *Enterococcus* and drive evolution of exopolysaccharide synthesis genes.

Co-authors: Clark Hendrickson, Cyril Samillano, and Katrine Whiteson

ABSTRACT

Phages infecting the diverse bacteria that colonize humans remain understudied but could be a valuable resource for treating antibiotic-resistant infections. *Enterococcus* is an opportunistic pathogen with rising rates of antibiotic resistance, but few phages infecting *Enterococcus* have been isolated and studied. Here, we characterize the *Brockvirinae* sub-family (formerly *Spounavirinae*) of phages that infect *Enterococcus* and present eight new phages within the sub-family. Genomic characterization reveals *Brockvirinae* phages represent two genera with distinct host range patterns. *In vitro* experimental evolution shows that as Brockvirinae phages co-evolve with their hosts, they select for *Enterococcus* with mutations in exopolysaccharide synthesis genes. Further, by searching the SRA, we show that these phages found globally in human and animal microbiomes. Characterizing the host range and molecular evolution of phages could lead to more efficient strategies for using phages therapeutically against antibiotic resistant bacteria.

INTRODUCTION

Bacteriophages (phages) are ubiquitous in the environment and important members of microbial communities. In the years following their discovery in 1915, phages were exploited to treat bacterial infections, which is referred to as phage therapy. The discovery of penicillin led to the abandonment of phage therapy in most of the western world, however the epidemic of antibiotic resistant bacterial infections has caused

renewed interest in phage therapy. Phages are much more host-specific than antibiotics, therefore a diverse collection of phages is required. This is a major hurdle for phage therapy to overcome because there is a lack of well-characterized phages that infect human pathogens.

Enterococcus is a gram-positive bacterium that is commonly found in low abundance in the gut of humans and animals (François Lebreton et al. 2017). The species *Enterococcus faecalis* and *Enterococcus faecium* are the most common species associated with humans. *Enterococcus* is an opportunistic pathogen and that has emerged as a major health crisis. It is considered a high-priority pathogen by the WHO due to high rates of antibiotic resistance and adaptation to the hospital environment (Raven et al. 2016; Lawe-Davies and Bennett 2017). The National Healthcare Safety Network reported that of hospital acquired *Enterococcus* infections, about 80% of *E. faecium,* and 8% of E. *faecalis* isolates demonstrated Vancomycin resistance. Phage therapy could offer an alternative approach to treating antibiotic resistant infections, and *Enterococcus* phages have shown promise as therapeutics *in vitro* (Khalifa et al. 2015b, 2018). However, few phages infecting *Enterococcus* have been isolated and characterized.

The *Enterococcus* phage family *Herelleviridae* (formerly the sub-family *Spounaviridae*) are tailed-phages with Myoviridae morphology and double-stranded DNA genomes between 127-157 kb in length (Klumpp et al. 2010; Barylski et al. 2018). *Herelleviridae* phages exclusively infect hosts in the phylum *Firmicutes*. *Enterococcus*-infecting phages within the *Herelleviridae* family are classified in the subfamily *Brockvirinae* and have been noted for their broad host ranges and potential as therapeutics (Khalifa et al. 2015a, 2016; Gelman et al. 2018).

Here, we present the genomes of eight newly isolated *Enterococcus* phages in the *Brockvirinae* subfamily. These represent two genera with distinct host ranges for *Enterococcus faecalis* and *faecium*. Like other members of *Brockvirinae*, these phages have a broad host range within their host genus. Further, we show that i*n vitro* growth with susceptible *Enterococcus* hosts results in predictable evolutionary outcomes, with the phages evolving capsid and tail fiber genes, and the *Enterococcus* hosts evolving exopolysaccharide synthesis genes. Understanding the evolutionary pressure these phages exert on exopolysaccharide synthesis genes could provide insight into combined therapies with antibiotics that also target the cell wall.

RESULTS

Genomics of new *Enterococcus* phages in the *Brockvirinae* sub-family

Eight novel *Enterococcus* phages in the *Brockvirinae* sub-family were isolated from sewage and their genomes sequenced (**Figure 4.1**). Average nucleotide identity (ANI) of core genes clearly divides the phages into two groups with ~95% ANI within each group and ~74 % ANI between the groups. Phages vB_OCPT_Car, vB_OCPT_Carl, and vB_OCPT_Bob fit into the *Kochikohdavirus* genus, and phages EfV12-phi1, vB_OCPT_Bop, vB_OCPT_Bill, vB_OCPT_Ben, vB_OCPT_Tex, and vB_OCPT_CCS1 fit into an Unassigned second genus, which we propose to call *Wandervirus*. Further, phages in the *Wandervirus* genus are split between two groups based on core genome nucleotide identity (97 % ANI within groups and ~94 % ANI between groups) and accessory genome content.

Brockvirinae phage genomes encode around 210 ORFs that are divided into two opposite facing blocks. The first block of genes encode short hypothetical ORFs (average 450 bp) of unknown function. These ORFs are often shared among phages within the phage

genus, and not shared between the two genera. The second block of ORFs is twice as long on average (900 bp) and encodes recognizable structural genes, lysins, and genes involved in genome replication. Most of these ORFs are conserved among all eleven *Brockvirinae* phages. Between these two gene blocks is a cluster of tRNA genes. Phages in the *Kochikohdavirus* genus carry 24 tRNAs on average while phages in the *Wandervirus* genus carry 7 tRNAs on average (F**igure S4.1**).

Brockvirinae phages are found in phage cocktails

To assess the environmental distribution of *Brockvirinae* phages, we queried a representative genome from both genera against 67,429 publicly available metagenomes in NCBI's SRA (**Table 4.1**) (Levi et al. 2018). Metagenomes with positive hits were downloaded and aligned to representative *Brockvirinae* genomes to ensure most of the genome was covered. *Brockvirinae* phages were found to be globally distributed in fecal metagenomes. Sequences matching *Brockvirinae* genomes were found in eight SRA projects from the United States, Europe, the Middle East, and Asia. Matching sequences were also found in non-human fecal metagenomes from condors, pigs, and bats. *Brockvirinae* phages were also found to be highly abundant in two phage cocktails from the Eliava Institute designed to treat intestinal issues. The first phage cocktail is the Intestiphage cocktail, which contains an isogenic *Brockvirinae* phage in the *Kochikohdavirus* genus. The second phage cocktail is the PYO phage cocktail developed by the Eliava institute.

Brockvirinae phages have broad host ranges

The host range for each of the nine novel *Brockvirinae* phages was tested against a collection of 36 *E. faecalis* strains and 29 *E. faecium* strains by drop assay. *Brockvirinae* phages demonstrated broad lytic activity within the two *Enterococcus* species (**Figure 4.2**).

No lysis was seen against any *Streptococcus* strains (data not shown). In general, *Kochikohdaviruses* infected *E. faecalis* strains but not *E. faecium*. Phages within *Wandervirus* A generally infected both *E. faecium* and *E. faecalis* strains, while phages belonging to *Wandervirus B* infected mostly *E. faecium* strains. These patterns indicate that genetic similarity among *Brockvirinae* phages is a good indicator of potential *Enterococcus* host species. Conversely, knowing the species of *Enterococcus* would provide insight into susceptibility to *Brockvirinae* phages. However, at the strain level, neither genetic similarity nor accessory genome content were predictive of susceptibility to *Brockvirinae* phages.

Brockvirinae phages drive evolution of *Enterococcus* exopolysaccharide synthesis genes

To understand the evolutionary pressures exerted between *Brockvirinae* phages and their *Enterococcus* hosts, pairs of bacteria and phage were experimentally coevolved *in vitro*, followed by whole genome sequencing. Co-evolution was performed by growing bacteria and phage together in semi-continuous liquid culture for four weeks. Cultures were started at an initial MOI of 0.01 so that *Enterococcus* were not completely lysed immediately. *Brockvirinae* phages evolved mutations in the same three genes regardless of host strain (**Table S4.4**). These genes encoded a "tail fiber gene", a "capsid and scaffold gene", and one ORF of unknown function. The ORF of unknown function contains a predicted ATPase domain and is homologous to genes in a wide range of phages and bacteria. The consistent mutations in the tail fiber gene and capsid and scaffold gene indicate that mutations in structural genes are the primary route for adapting to hosts in these conditions.

Enterococcus strains grown with *Brockvirinae* phages consistently evolved mutations in genes involved in exopolysaccharide synthesis (**Figure 4.3, Table S4.3**). All

three *E. faecalis* strains acquired mutations in genes in the Epa exopolysaccharide synthesis locus. Of the approximately nineteen genes in the Epa locus, three genes consistently developed point mutations and nonsense mutations in *E. faecalis* B3286, *E. faecalis* TX2137, and *E. faecalis* Yi6. Host control cultures lacking phage never acquired mutations in these genes. In contrast to *E. faecalis*, when *E. faecium* TX1330 coevolved with *Brockvirinae* phages, the Yqw exopolysaccharide synthesis locus was consistently mutated. However, mutations in the Yqw locus mutations also occurred in the host control cultures. However, host control cultures of *E. faecium* also showed signs of prophage induction based on an increase in sequencing coverage of a predicted prophage, therefore it is unclear if mutations in the Yqw locus can be attributed to phage evolutionary pressure. Mutations in other genes were seen, but not with the same frequency or consistency as genes involved in exopolysaccharide synthesis (**Table S4.3**). Therefore, mutations in capsule synthesis genes are a consistent feature of E*nterococcus* coevolution with *Brockvirinae* phages.

DISCUSSION

Here, we present a characterization of two genera of *Brockvirinae* phages infecting *Enterococcus*. These two genera display distinct infectivity patterns for *E. faecium* and *E. faecalis*. *Brockvirinae* phages were found to be globally distributed in fecal metagenomes of humans and several animals. Experimental coevolution of these phages with their host bacteria consistently results in bacterial mutations in exopolysaccharide synthesis loci and phage mutations in structural genes.

The *Brockvirinae* family contains phages that infect *Firmicutes*. While the genome nucleotide identity and amino acid identities are extremely low when comparing the *Spounaviridae* genera, they share a common morphology and genome organization

(Barylski et al. 2018; Hejnowicz et al. 2012). Phages are often thought to partake in rampant horizontal gene transfer resulting in a high level of genome mosaicism that clouds phylogenetic relationships. However, little evidence of mosaicism has been previously seen in *Herelleviridae* phages, and genome nucleotide identity tracks with host range down to the phage sub-genus level (Barylski et al. 2018; Bolduc et al. 2017). Similarly, we saw little evidence of mosaicism in *Enterococcus*-infecting *Herelleviridae* phages in either gene content or nucleotide identity.

Herelleviridae phages have been previously observed to have broad host ranges within the genus they infect. Likewise, we saw *Enterococcus*-infecting *Brockvirinae* phages display broad host ranges for *E. faecium* and *E. faecalis*. The lysin from an *Enterococcus*-infecting *Brockvirinae* phage, EfV12-phi1, has been shown to lyse different genera, including *Staphylococcus* and *Streptococcus* (Yoong et al. 2004a). However, no evidence exists that any *Spounaviridae* phages can infect genera other than their host.

Herelleviridae phages were confidently found in twelve sequencing projects out of the thousands of human gut metagenomes from the SRA. Since *Enterococcus* is present in the gut of most people, this suggests that either *Herelleviridae* phages are uncommon members of the human gut, or they are usually at a low enough abundance that they would not be seen with a normal depth of shotgun sequencing. We are inclined to the second explanation given the high frequency which we have isolated *Herelleviridae* phages from sewage. Even one of the most abundant phages in the human microbiome, crAssphage infecting *Bacteroides*, is only found at 1% read abundance in human metagenomes, so a phage infecting a minority community member such as *Enterococcus* would be much less abundant.

Structural gene mutations were a consistent feature seen in phage evolution. Two phages in the *Wandervirus* genus were seen to mutate the same genes when evolving with different *E. faecium* and *E. faecalis* hosts. These genes were primarily structural genes encoding the capsid and tail fibers and likely are involved in the initial binding of phage to the bacterial surface. A common way bacteria evolve resistance to phage infection is to prevent binding to the cell surface, and mutations in genes such as the tail fiber can overcome that resistance. Phage structural gene mutations are commonly seen to mutate when phages co-evolve *in vitro* with their host (Wandro et al. 2019; Uchiyama et al. 2011). Tail fiber genes mediate host binding mutations can change binding affinity to certain hosts (Wandro et al. 2019; Uchiyama et al. 2011; Perry et al. 2015). The binding target of *Brockvirinae* phages has not been characterized, but related *Spounaviridae* phages have been shown to bind teichoic acids in the cell wall (ref).

When under selective pressure from *Brockvirinae* phages, *Enterococcus* evolved mutations primarily in exopolysaccharide synthesis genes. These mutations suggest resistance is evolving by preventing phage recognition and initial binding. *E. faecium* and *E. faecalis* both contain the highly conserved Epa capsule synthesis locus, in which mutations were observed consistently for *E. faecalis* strains. Mutations in the Epa locus have been observed previously during coevolution with *Brockvirinae* phages and other phages impaired *Enterococcus* host colonization and increased antibiotic sensitivity (Chatterjee et al. 2019). There is also a second locus that differs between the two species: CPS in *E. faecalis* and Yqw in *E. faecium*. It is in this Yqw locus that mutations were observed in *E. faecium* TX1330. Various mutations in a single gene of this locus have been previously seen during coevolution in the same phage-host pair (Wandro et al. 2019). Since mutations in

Yqw locus genes also occurred during *E. faecium* TX1330 host control cultures lacking phage, we cannot attribute these mutations to phage evolutionary pressure. However, prophage induction was observed in all *E. faecium* TX1330 cultures. Together, this suggests that mutations in exopolysaccharide synthesis genes may be a consistent feature of *Enterococcus* evolution to phage in general.

The broad host range of *Brockvirinae* phages and predictable outcomes of coevolution with their hosts make them ideal candidates for use in phage therapy to treat *Enterococcus* infections. Multiple commercial phage therapy cocktails already include *Brockvirinae* phages. Although *Enterococcus* are seen to evolve resistance to infection, including a diverse set of phages in a phage cocktail could lessen that effect (M. Yen, Cairns, and Camilli 2017b; Nale et al. 2018a). Further, there may be trade-offs so that mutating resistance to phage results in a less fit *Enterococcus* (Chatterjee et al. 2019). Phage therapy is a promising avenue of research for treating antibiotic resistant infections, but more work needs to be put into isolating and characterizing collections of phages that infect important pathogens.

MATERIALS AND METHODS

Bacteria and phage strains and growth conditions

Enterococcus isolates were either ordered from BEI or obtained from UC San Diego clinical microbiology laboratory (**Table S4.2**). *Enterococcus* was grown statically at 37 °C in brain heart infusion (BHI) media in all experiments. Phage EfV12-phi1 was ordered from Felix d'Herelle Reference Center for Bacterial Viruses (HER# 339). All other phages were isolated from sewage (**Table S4.1**).

Phage isolation propagation and storage

Phages were isolated from sewage using three rounds of plaque assays. Raw sewage influent was collected from wastewater treatment plants in Redwood Shores and Escondido, California. Sewage was stored at 4 °C and used for phage isolation for several months. Sewage was centrifuged for 10 minutes at 10,000 g to remove particulates and the supernatant was used in plaque assays with various strains of *Enterococcus*. 100 ul sewage supernatant was added to 100ul exponentially growing *Enterococcus* in BHI media and incubated at 37 °C for 15 minutes. 5 mL of warm BHI containing 0.3 % UltraPure Low Melting Point Agarose (ThermoFisher #16520050) was then added and the mixture poured on a BHI agar plate and incubated overnight at 37 °C. The next day, plates were examined for plaques and any plaques are picked with a pipette tip and suspended in 50 ul SM buffer. Picked plaques underwent two more rounds of plaque assays in the same manner to ensure purity of the phage isolate. Pure phages were propagated by performing plaque assay to create a plate displaying webbed lysis that was then flooded with 3 mL SM buffer and incubated for 1 hour. The SM buffer was then collected and centrifuged at 10,000 g for 10 minutes. For long term storage, phages were stored at -80 °C with 25% glycerol.

Genomic sequencing

DNA was extracted from *Enterococcus* and phage using Quick-DNA Microprep Kit (Zymo #D3020). Before *Enterococcus* DNA extraction, lysozyme was added to lysis buffer at a concentration of 100ug/ml and incubated at 37 °C for 30 minutes. For DNA extraction from coevolution cultures containing both bacteria and phage, the extractions were performed without lysozyme. Libraries were prepared using a scaled down reactions with the Illumina Nextera enzyme (Baym et al. 2015). Paired-end sequencing with a 75 bp read

length was performed on the Illumina NextSeq using the Mid Output v2 reagents. Approximately 2.5 million reads were obtained for each sample.

Genomic characterization

Phage and *Enterococcus* genomes were assembled *de novo* using the SPAdes assembler (Bankevich et al. 2012). Core genomes were determined and aligned using Anvio (Eren et al. 2015). Genomes were manually examined using Geneious.

SRA search

All metagenomes in the SRA were searched for *Brockvirinae Enterococcus* phages using the Searching SRA tool using the core genome as the query sequence (Levi et al. 2018). Briefly, the Searching SRA tool searches for the query sequence in all 111,156 metagenomes currently on the SRA by subsampling 100,000 sequences from each metagenome. From the metagenome hit list, we selected only metagenomes where the average read length matching our query was over 50 bp.

Host range assay

Phage host ranges were tested using a spot assay. 5 mL of warm BHI containing 0.3 % UltraPure Low Melting Point Agarose (ThermoFisher #16520050) was added to 100 ul of exponentially growing *Enterococcus* and poured on a BHI agar plate. After allowing the agarose to solidify for approximately 30 minutes, 5 ul droplets of each concentrated phage was spotted on top of the agarose. As a negative control, SM buffer was spotted in the same fashion. Spots were allowed to dry for 30 minutes and then the plates were incubated overnight at 37 °C. The next day, each spot was checked for clearing.

Coevolution of *Enterococcus* and phage

Pairs of *Enterococcus* strains and phage isolates were co-evolved in liquid media for 28 days with once daily dilution. To start the culture, phage was added to exponentially growing *Enterococcus* at an MOI of approximately 0.01 in 200 ul BHI liquid media in a 96-well plate. Plate was grown statically for 24 hours at 37 °C, then 10 ul of each well was diluted into 190 ul fresh BHI media in a new 96-well plate. This process was repeated for 28 days. At the end of the experiment, 150 ul of the final cultures were pipetted into a new 96-well plate and 150 ul of 50 % glycerol was added and the plate was stored at -80 °C prior to DNA extraction.

FIGURES AND TABLES

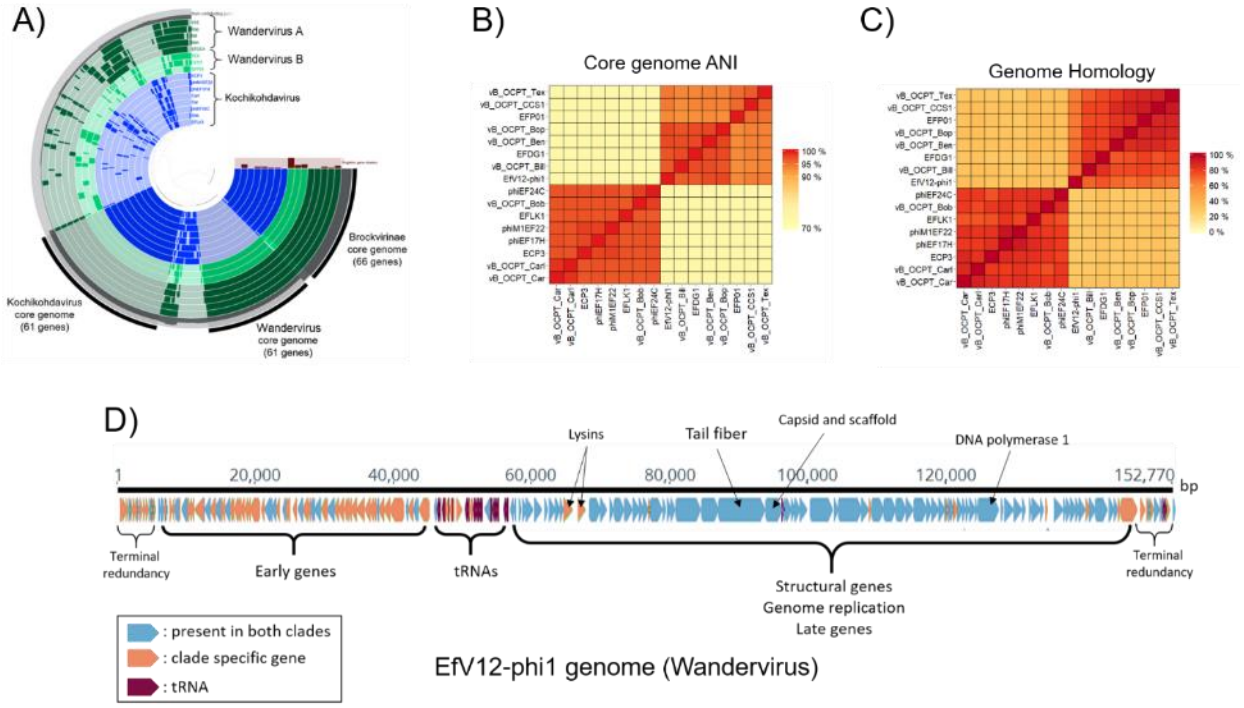


Figure 4.1. Genomics of *Enterococcus Brockvirinae* phages. **A)** Known *Brockvirinae* *Enterococcus* phages are divided into two clades. **B)** Core genome average nucleotide identity of all *Brockvirinae* phages. **C)** Overall genome homology of *Brockvirinae* phages. **D)** Genome of phage EfV12-phi1 showing some genes in *Brockvirinae* phages are conserved at the genus level and some are present in both genera.

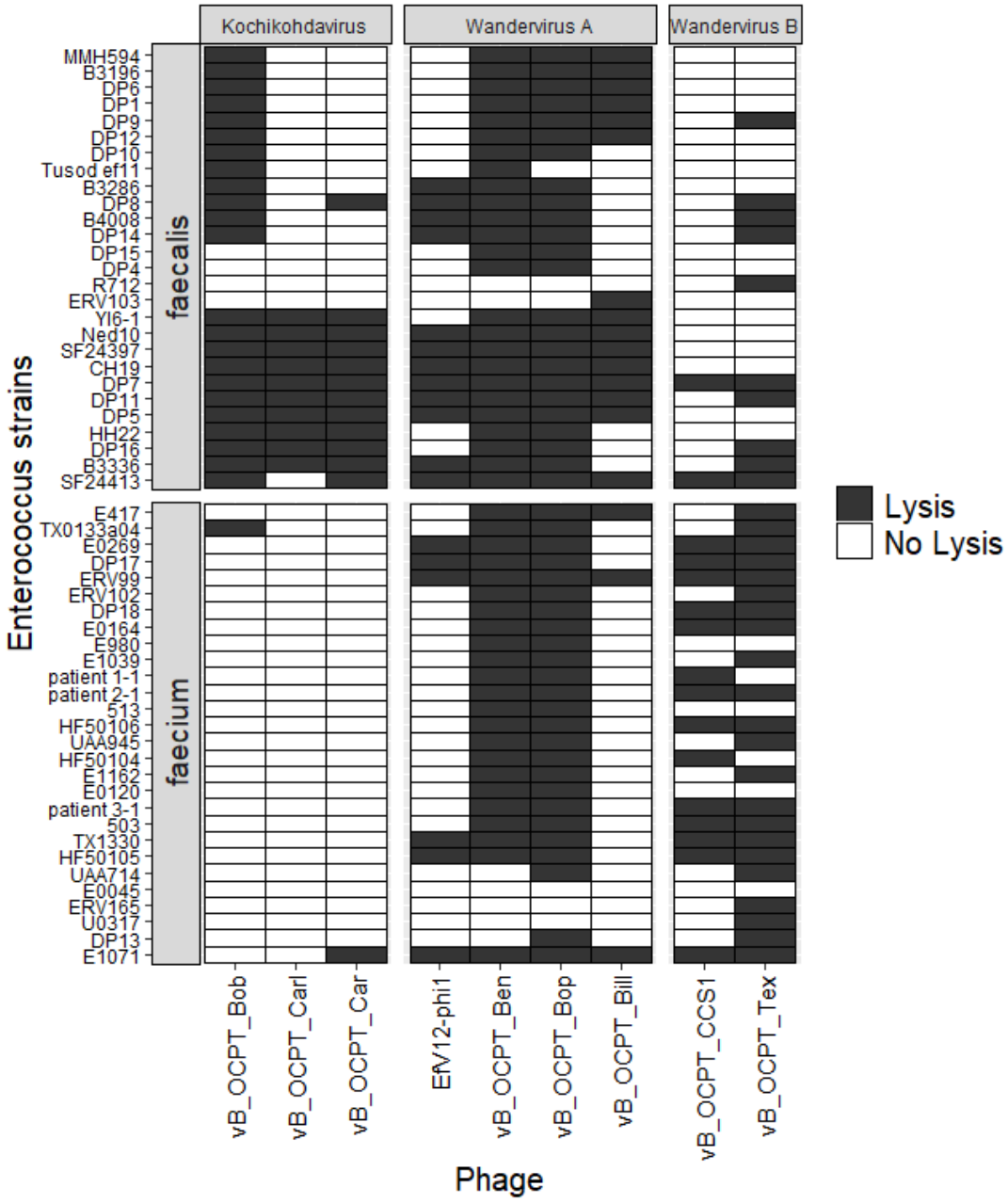


Figure 4.2. Host range of *Brockvirinae Enterococcus* phages. Host range was determined by drop assay and visually scored. Partial clearings were counted as lysis.

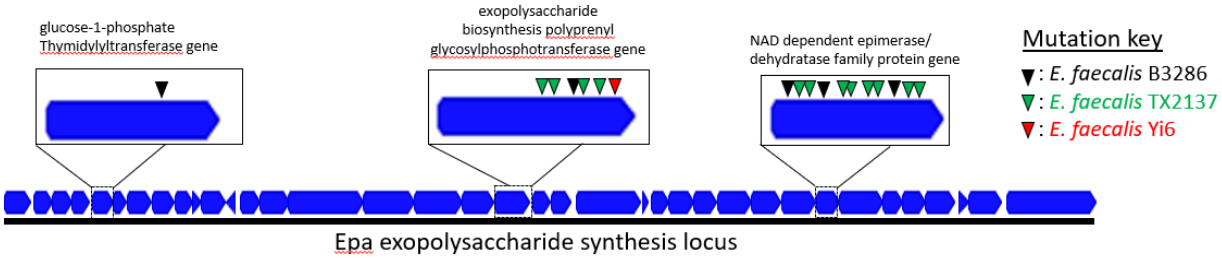


Figure 4.3. *E. faecalis* strains evolve mutations in Epa locus genes to resist phage. The genes comprising the Epa locus of *E. faecalis* is shown. Black, green, and red ticks represent the locations of non-synonymous mutations observed in *E. faecalis* B3286, TX2137, and Yi6 respectively as they coevolve with *Brockvirinae* phages. Detailed information about these mutations can be found in **Table S4.3**.

Table 4.1. *Brockvirinae Enterococcus* phages from the Sequence Read Archive.

SRA	Title	Location	Sample type
SRP077952	The INTESTI bacteriophage cocktail genome sequencing and assembly	Georgia	
PRJEB23244	PYO phage cocktail	Georgia	Phage cocktail
ERP017091	The gut microbiome in Crohn's disease and modulation by exclusive enteral nutrition	Guangdong, China	human fecal
ERP006678	Gut and Oral Microbiome Dysbiosis in Rheumatoid Arthritis	Beijing China	human fecal
SRP071229	Gymnogyps californianus microbiome raw sequence reads	Los Alamos National Laboratory	California condor fecal
ERP006046	Virus_Discovery_for_Vietnam_Initiative_on_Zoonotic_Infections__VIZIONS_	Vietnam	viral metagenome
ERP001956	Diagnostic Metagenomics: A Culture-Independent Approach to the Investigation of Bacterial Infections	Germany	human fecal
SRP051511	New York City MTA subway samples Metagenome	New York City	subway samples
ERP012929	Towards personalized nutrition by prediction of glycemic responses	Israel	human fecal
SRP040146	Clostridium difficile FMT	Broad institute, Massachusetts	human fecal
SRP115494	Longitudinal Multi'omics of the Human Microbiome in Inflammatory Bowel Disease		
SRP099123	Metagenomic analysis of gut microbiota in sows and piglets	Freie University of Berlin	Pig fecal

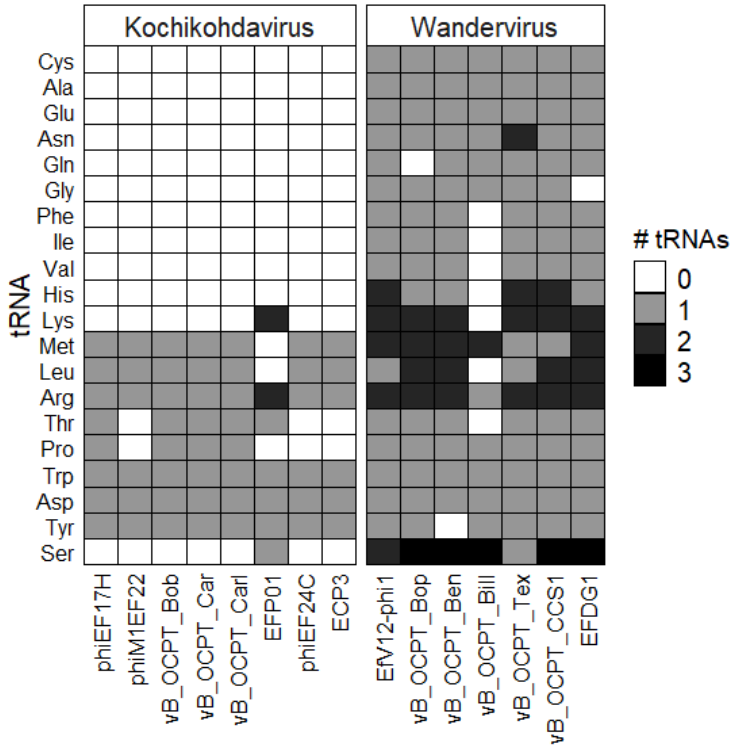


Figure S4.1. tRNAs in phage genomes.

Table S4.1. *Brockvirinae* phage information.

Phage	genus	isolation source	Genbank
vB_OCPT_Bob	Kochikohdavirus	Escondido sewage	
vB_OCPT_Car	Kochikohdavirus	Escondido sewage	
vB_OCPT_Carl	Kochikohdavirus	Escondido sewage	
EfV12-phi1	Wandervirus A	Canadian sewage	MH880817.1
vB_OCPT_Ben	Wandervirus A	Escondido sewage	
vB_OCPT_Bop	Wandervirus A	Escondido sewage	
vB_OCPT_Bill	Wandervirus A	Escondido sewage	
vB_OCPT_CCS1	Wandervirus B	Escondido sewage	
vB_OCPT_Tex	Wandervirus B	Escondido sewage	

Table S4.2. *Enterococcus* strain information.

Species	Strain	Genbank	Source
Enterococcus faecalis	B3319	GCA_000396325.1	HMP
Enterococcus faecalis	B3196	GCA_000396345.1	HMP
Enterococcus faecalis	B3286	GCA_000396365.1	HMP
Enterococcus faecalis	B3336	GCA_000396385.1	HMP
Enterococcus faecalis	B4008	GCA_000396405.1	HMP
Enterococcus faecalis	CH19	GCA_000394255.1	HMP
Enterococcus faecalis	HH22	GCA_000394775.1	HMP
Enterococcus faecalis	MMH594	GCA_000394795.1	HMP
Enterococcus faecalis	Ned10	GCA_000394875.1	HMP
Enterococcus faecalis	R712	GCA_000163815.1	HMP
Enterococcus faecalis	S613	GCA_000163795.1	HMP
Enterococcus faecalis	SF105	GCA_000394895.1	HMP
Enterococcus faecalis	SF24397	GCA_000394075.1	HMP
Enterococcus faecalis	SF24413	GCA_000394095.1	HMP
Enterococcus faecalis	SF28073	GCA_000394195.1	HMP
Enterococcus faecalis	Tusod ef11	GCA_000175015.1	HMP
Enterococcus faecalis	TX1322	GCA_000159275.1	HMP
Enterococcus faecalis	TX2137	GCA_000147595.1	HMP
Enterococcus faecalis	V587	GCA_000394175.1	HMP
Enterococcus faecalis	YI6-1	GCA_000395095.1	HMP
Enterococcus faecalis	DP1		David Pride hospital isolate
Enterococcus faecalis	DP2		David Pride hospital isolate
Enterococcus faecalis	DP12		David Pride hospital isolate
Enterococcus faecalis	DP14		David Pride hospital isolate
Enterococcus faecalis	DP15		David Pride hospital isolate
Enterococcus faecalis	DP16		David Pride hospital isolate
Enterococcus faecalis	ERV103	GCA_000294005.2	HMP
Enterococcus faecalis	DP3		David Pride hospital isolate
Enterococcus faecalis	DP4		David Pride hospital isolate
Enterococcus faecalis	DP5		David Pride hospital isolate
Enterococcus faecalis	DP6		David Pride hospital isolate
Enterococcus faecalis	DP7		David Pride hospital isolate
Enterococcus faecalis	DP8		David Pride hospital isolate
Enterococcus faecalis	DP9		David Pride hospital isolate
Enterococcus faecalis	DP10		David Pride hospital isolate
Enterococcus faecalis	DP11		David Pride hospital isolate
Enterococcus faecium	E0120	GCA_000321485.1	HMP
Enterococcus faecium	E0164	GCA_000321505.1	HMP

Enterococcus faecium	E0269	GCA_000321525.1	HMP
Enterococcus faecium	ERV102	GCA_000295355.2	HMP
Enterococcus faecium	ERV99	GCA_000295175.2	HMP
Enterococcus faecium	HF50104	GCA_000396685.1	HMP
Enterococcus faecium	HF50105	GCA_000396705.1	HMP
Enterococcus faecium	HF50106	GCA_000396725.1	HMP
Enterococcus faecium	patient 1-1	GCA_000394755.1	HMP
Enterococcus faecium	patient 2-1	GCA_000394655.1	HMP
Enterococcus faecium	patient 3-1		HMP
Enterococcus faecium	TX0133a04	GCA_000147235.1	HMP
Enterococcus faecium	TX1330	GCA_003583905.1	HMP
Enterococcus faecium	UAA714	GCA_000395445.1	HMP
Enterococcus faecium	503	GCA_000295055.2	HMP
Enterococcus faecium	513	GCA_000295575.2	HMP
Enterococcus faecium	E0045	GCA_000321465.1	HMP
Enterococcus faecium	E1071	GCA_000172655.1	HMP
Enterococcus faecium	E1039	GCA_000174935.1	HMP
Enterococcus faecium	UAA945	GCA_000396845.1	HMP
Enterococcus faecium	U0317	GCA_000172915.1	HMP
Enterococcus faecium	E417	GCA_000295415.2	HMP
Enterococcus faecium	DP13		David Pride hospital isolate
Enterococcus faecium	DP17		David Pride hospital isolate
Enterococcus faecium	DP18		David Pride hospital isolate
Enterococcus faecium	DP19		David Pride hospital isolate
Enterococcus faecium	E1162	GCA_000172675.1	HMP
Enterococcus faecium	ERV165	GCA_000295235.2	HMP
Enterococcus faecium	E980	GCA_0001726151	HMP
Enterococcus avium	DP0		David Pride hospital isolate

Table S4.3. *Enterococcus* mutations. List of all mutations that occurred in *Enterococcus* genomes during coevolution with phage for 28 days.

Species	Strain	Coevolving Phage	Rep	Freq.	Locus	Gene name	Type
E. faecalis	B3286	vB_OCPT_Bop	1	1.00	Epa	NAD-dependent epimerase/dehydratase	SNP
E. faecalis	B3286	vB_OCPT_Bop	1	1.00	Epa	glucose-1-phosphate thymidylyltransferase	SNP
E. faecalis	B3286	vB_OCPT_Bop	2	1.00		endonuclease III	SNP
E. faecalis	B3286	vB_OCPT_Bop	2	1.00	Epa	NAD-dependent epimerase/dehydratase	SNP
E. faecalis	B3286	EfV12-phi1	1	1.00	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	B3286	EfV12-phi1	1	0.61	Epa	NAD-dependent epimerase/dehydratase	SNP
E. faecalis	B3286	EfV12-phi1	1	1.00		ATP-dependent Clp protease ATP-binding subunit ClpE	SNP
E. faecium	TX1330	vB_OCPT_Ben	1	1.00		DNA gyrase subunit A (EC 5.99.1.3)	SNP
E. faecium	TX1330	vB_OCPT_Ben	2	1.00		DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	SNP
E. faecium	TX1330	vB_OCPT_Ben	2	0.94	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Ben	3	0.50	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Ben	3	1.00		DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)	SNP
E. faecium	TX1330	vB_OCPT_Bill	1	1.00	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Bill	2	0.22	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Bob	1	0.26	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecium	TX1330	vB_OCPT_Bob	1	0.15		Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	SNP
E. faecium	TX1330	vB_OCPT_Bob	2	0.22		Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	SNP
E. faecium	TX1330	vB_OCPT_Bob	3	0.11		Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	SNP
E. faecium	TX1330	vB_OCPT_Bob	3	1.00		Two-component transcriptional response regulator, LuxR family	SNP
E. faecium	TX1330	vB_OCPT_Bop	1	1.00		UDP-N-acetylglucosamine 4,6-dehydratase (EC 4.2.1.135)	SNP
E. faecium	TX1330	vB_OCPT_Bop	1	1.00	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Bop	2	1.00	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	vB_OCPT_Bop	2	0.30		Neopullulanase (EC 3.2.1.135)	SNP
E. faecium	TX1330	vB_OCPT_Bop	2	0.29		hypothetical protein	SNP
E. faecium	TX1330	vB_OCPT_Car	1	0.41		Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	SNP
E. faecium	TX1330	vB_OCPT_Car	1	0.26		Bacterial ribosome SSU maturation protein RimP	SNP
E. faecium	TX1330	vB_OCPT_Carl	2	1.00	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecium	TX1330	vB_OCPT_Carl	2	0.26		Xanthine/uracil/thiamine/ascorbate permease family protein	SNP
E. faecium	TX1330	vB_OCPT_Carl	2	0.21		Oxidoreductase, short-chain dehydrogenase/reductase family	SNP
E. faecium	TX1330	vB_OCPT_Carl	2	0.20		Excinuclease ABC subunit A paralog of unknown function	SNP
E. faecium	TX1330	vB_OCPT_Carl	3	0.34	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecium	TX1330	No phage	1	1.00	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecium	TX1330	No phage	1	0.27		Transcriptional regulator, repressor of the glutamine synthetase, MerR family	SNP
E. faecium	TX1330	No phage	1	0.26		Two-component sensor kinase SA14-24	SNP

E. faecium	TX1330	No phage	1	0.24		Cyclic-di-AMP phosphodiesterase GdpP	SNP
E. faecium	TX1330	No phage	1	0.24		Ribonuclease PH (EC 2.7.7.56)	SNP
E. faecium	TX1330	No phage	1	0.23		Two-component system YycFG regulatory protein YycH	SNP
E. faecium	TX1330	No phage	2	1.00	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecium	TX1330	No phage	2	0.23		Teichoic acid glycosylation protein	SNP
E. faecium	TX1330	No phage	3	0.70	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	No phage	3	0.30	Yqw	Tyrosine-protein kinase EpsD (EC 2.7.10.2)	SNP
E. faecium	TX1330	No phage	4	0.24		ABC transporter, permease protein YckA (cluster 3, basic aa/glutamine/opines)	SNP
E. faecium	TX1330	V12	1	1.00	Yqw	Tyrosine-protein kinase transmembrane modulator EpsC	SNP
E. faecalis	TX2137	Bop	1	0.63	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	1	0.31		UDP-N-acetylglucosamine 1-carboxyvinyltransferase	SNP
E. faecalis	TX2137	Bop	1	0.12	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	2	1.00		phosphocarrier protein HPr	SNP
E. faecalis	TX2137	Bop	2	1.00	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	3	0.90	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	Bop	3	0.89	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	Bop	4	0.25	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	Bop	4	0.25	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	Bop	4	0.20	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	5	0.17		UDP-N-acetylglucosamine 1-carboxyvinyltransferase	SNP
E. faecalis	TX2137	Bop	6	0.25	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	6	0.25	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	6	0.17	Epa	NAD dependent epimerase/dehydratase family protein	DEL
E. faecalis	TX2137	Bop	7	0.82	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	7	0.17	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	8	1.00	Epa	NAD dependent epimerase/dehydratase family protein	SNP
E. faecalis	TX2137	Bop	8	0.59		UDP-N-acetylglucosamine 1-carboxyvinyltransferase	SNP
E. faecalis	TX2137	V12	1	0.71	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	V12	1	0.71	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	TX2137	V12	2	0.69		DNA ligase (NAD+)	SNP
E. faecalis	TX2137	V12	2	0.55	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	DEL
E. faecalis	TX2137	V12	2	0.31	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	Yi6	Bop	1	1.00	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SUB
E. faecalis	Yi6	Bop	1	1.00	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	Yi6	Bop	1	1.00	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP
E. faecalis	Yi6	Bop	1	0.19	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	SNP

E. faecalis	Yi6	Bop	2	1.00		UTP-glucose-1-phosphate uridylyltransferase	DEL
E. faecalis	Yi6	Bop	2	1.00		UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	SNP
E. faecalis	Yi6	Bop	3	1.00		isoleucyl-tRNA synthetase	DEL
E. faecalis	Yi6	Bop	3	0.87	Epa	exopolysaccharide biosynthesis polyprenyl glycosylphosphotransferase	DEL
E. faecalis	Yi6	V12	1	0.53		30S ribosomal protein S7	SNP
E. faecalis	Yi6	V12	2	0.50		DNA-directed RNA polymerase subunit alpha	SNP

Table S4.4. Phage mutations. List of all mutations that occurred in phage genomes during coevolution with *Enterococcus* for 28 days.

Phage	Rep	Host species	Host strain	Freq	Type	Gene
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
EfV12-phi1	1	E. faecalis	B3286	0.90	SNP	Phage recombination related exonuclease (EC 3.1.11.-)
EfV12-phi1	1	E. faecalis	B3286	1.00	DEL	Phage protein
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	B3286	0.65	SNP	Phage protein
EfV12-phi1	1	E. faecalis	B3286	0.91	SNP	Protein RtcB
EfV12-phi1	1	E. faecalis	B3286	0.40	SNP	Phage protein
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
EfV12-phi1	2	E. faecalis	B3286	0.44	SNP	hypothetical protein
EfV12-phi1	2	E. faecalis	B3286	0.84	SNP	Phage capsid and scaffold
EfV12-phi1	2	E. faecalis	B3286	0.51	SNP	Phage protein
EfV12-phi1	2	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
EfV12-phi1	2	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	B3286	0.22	SNP	Protein RtcB
vB_OCPT_Bop	3	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	3	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	B3286	0.46	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	4	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	4	E. faecalis	B3286	0.24	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	B3286	0.53	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	B3286	1.00	SNP	hypothetical protein
vB_OCPT_Bop	4	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	5	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	5	E. faecalis	B3286	0.80	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	B3286	0.81	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	B3286	0.80	SNP	Serine/threonine protein phosphatase (EC 3.1.3.16)
vB_OCPT_Bop	5	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant

vB_OCPT_Bop	6	E. faecalis	B3286	0.19	SNP	hypothetical protein
vB_OCPT_Bop	6	E. faecalis	B3286	0.11	SNP	hypothetical protein
vB_OCPT_Bop	6	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	6	E. faecalis	B3286	0.18	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	6	E. faecalis	B3286	0.56	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	B3286	0.23	SNP	Nicotinamide-nucleotide adenylyltransferase, NadR family (EC 2.7.7.1) / Ribosylnicotinamide kinase (EC 2.7.1.22)
vB_OCPT_Bop	6	E. faecalis	B3286	0.13	SNP	Phage protein
vB_OCPT_Bop	6	E. faecalis	B3286	0.21	SNP	DNA helicase, phage-associated
vB_OCPT_Bop	6	E. faecalis	B3286	0.19	SNP	Nicotinamide-nucleotide adenylyltransferase, NadR family (EC 2.7.7.1) / Ribosylnicotinamide kinase (EC 2.7.1.22)
vB_OCPT_Bop	6	E. faecalis	B3286	0.15	SNP	Phage major tail sheath
vB_OCPT_Bop	6	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	6	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	6	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	B3286	0.25	SNP	Deoxyadenosine kinase (EC 2.7.1.76) / Deoxyguanosine kinase (EC 2.7.1.113)
vB_OCPT_Bop	7	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	0.22	SNP	Phage baseplate
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	INS	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	0.25	SNP	hypothetical protein
vB_OCPT_Bop	8	E. faecalis	B3286	0.22	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	B3286	0.18	SNP	hypothetical protein
vB_OCPT_Bop	8	E. faecalis	B3286	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	8	E. faecalis	B3286	0.22	SNP	NrdR-regulated deoxyribonucleotide transporter, PnuC-like
vB_OCPT_Bop	8	E. faecalis	B3286	0.24	SNP	Phage protein
EfV12-phi1	1	E. faecalis	TX2137	0.38	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	TX2137	0.40	SNP	Phage capsid and scaffold
EfV12-phi1	1	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
EfV12-phi1	2	E. faecalis	TX2137	0.58	SNP	hypothetical protein
EfV12-phi1	2	E. faecalis	TX2137	0.92	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
EfV12-phi1	3	E. faecalis	TX2137	0.80	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	TX2137	0.17	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	TX2137	0.39	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant

vB_OCPT_Bop	4	E. faecalis	TX2137	0.91	SNP	Phage capsid and scaffold
vB_OCPT_Bop	4	E. faecalis	TX2137	0.17	SNP	hypothetical protein
vB_OCPT_Bop	4	E. faecalis	TX2137	0.84	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	5	E. faecalis	TX2137	0.17	SNP	hypothetical protein
vB_OCPT_Bop	6	E. faecalis	TX2137	0.38	SNP	Ribonucleotide reductase of class III (anaerobic), activating protein (EC 1.97.1.4)
vB_OCPT_Bop	6	E. faecalis	TX2137	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	TX2137	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	TX2137	0.40	SNP	Phage protein
vB_OCPT_Bop	7	E. faecalis	TX2137	0.95	SNP	Phage capsid and scaffold
vB_OCPT_Bop	7	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	8	E. faecalis	TX2137	0.98	SNP	Phage capsid and scaffold
vB_OCPT_Bop	8	E. faecalis	TX2137	0.99	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	TX2137	0.36	SNP	Phage protein
vB_OCPT_Bop	8	E. faecalis	TX2137	0.11	SNP	Thioredoxin, phage-associated
vB_OCPT_Bop	8	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	9	E. faecalis	TX2137	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	9	E. faecalis	TX2137	0.57	SNP	Phage protein
vB_OCPT_Bop	9	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	10	E. faecalis	TX2137	0.66	SNP	Phage capsid and scaffold
vB_OCPT_Bop	10	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	10	E. faecalis	TX2137	0.97	SNP	Phage capsid and scaffold
vB_OCPT_Bop	11	E. faecalis	TX2137	0.84	SNP	Phage capsid and scaffold
vB_OCPT_Bop	11	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	11	E. faecalis	TX2137	0.18	SNP	Phage capsid and scaffold
vB_OCPT_Bop	12	E. faecalis	TX2137	0.21	SNP	DNA primase (EC 2.7.7.-) / DNA helicase (EC 3.6.1.-), phage-associated
vB_OCPT_Bop	12	E. faecalis	TX2137	0.97	SNP	Phage capsid and scaffold
vB_OCPT_Bop	12	E. faecalis	TX2137	0.16	SNP	Phage protein
vB_OCPT_Bop	12	E. faecalis	TX2137	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	12	E. faecalis	TX2137	0.11	SNP	Phage protein
vB_OCPT_Bop	13	E. faecalis	TX2137	0.25	SNP	Phage protein
vB_OCPT_Bop	13	E. faecalis	TX2137	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	13	E. faecalis	TX2137	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	13	E. faecalis	TX2137	0.38	SNP	Phage protein
vB_OCPT_Bop	1	E. faecalis	Yi6	0.54	SNP	Phage protein
vB_OCPT_Bop	1	E. faecalis	Yi6	0.38	SNP	Phage protein
vB_OCPT_Bop	1	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	2	E. faecalis	Yi6	0.98	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	2	E. faecalis	Yi6	1.00	SNP	Phage protein
vB_OCPT_Bop	2	E. faecalis	Yi6	0.96	SNP	Phage protein
vB_OCPT_Bop	2	E. faecalis	Yi6	1.00	SNP	Phage protein

vB_OCPT_Bop	2	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	2	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	Yi6	0.96	SNP	Phage capsid and scaffold
vB_OCPT_Bop	3	E. faecalis	Yi6	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	3	E. faecalis	Yi6	1.00	SNP	Phage protein
vB_OCPT_Bop	3	E. faecalis	Yi6	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	4	E. faecalis	Yi6	0.95	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	Yi6	0.12	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	Yi6	0.40	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	Yi6	0.79	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	4	E. faecalis	Yi6	0.99	SNP	Phage protein
vB_OCPT_Bop	4	E. faecalis	Yi6	1.00	SNP	Phage protein
vB_OCPT_Bop	5	E. faecalis	Yi6	0.64	SNP	Phage protein
vB_OCPT_Bop	5	E. faecalis	Yi6	0.32	SNP	Phage protein
vB_OCPT_Bop	5	E. faecalis	Yi6	0.13	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	0.16	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	0.49	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	0.76	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	5	E. faecalis	Yi6	0.82	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	0.24	SNP	Phage capsid and scaffold
vB_OCPT_Bop	5	E. faecalis	Yi6	0.10	SNP	hypothetical protein
vB_OCPT_Bop	6	E. faecalis	Yi6	0.82	SNP	Phage protein
vB_OCPT_Bop	6	E. faecalis	Yi6	0.13	SNP	Phage protein
vB_OCPT_Bop	6	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold
vB_OCPT_Bop	6	E. faecalis	Yi6	1.00	SNP	Phage protein
vB_OCPT_Bop	6	E. faecalis	Yi6	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	7	E. faecalis	Yi6	0.99	SNP	Phage protein
vB_OCPT_Bop	7	E. faecalis	Yi6	0.98	SNP	Phage capsid and scaffold
vB_OCPT_Bop	7	E. faecalis	Yi6	1.00	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	7	E. faecalis	Yi6	0.20	SNP	Phage protein
vB_OCPT_Bop	7	E. faecalis	Yi6	0.91	SNP	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46), phage variant
vB_OCPT_Bop	8	E. faecalis	Yi6	1.00	SNP	Phage capsid and scaffold

CHAPTER 5

Phage cocktails can prevent the evolution of phage-resistant *Enterococcus*

Co-authors: Clark Hendrickson, Cyril Samillano, and Katrine Whiteson

ABSTRACT

Antibiotic resistant *Enterococcus* infections are a major health crisis that requires the development of alternative therapies. Phage therapy could be an alternative to antibiotics and has shown promise in *in vitro* and in early trials. Phage therapy is often deployed as a cocktail of phages, but there is little understanding about how to combine phages in effective ways. Here we utilize a collection of 15 *Enterococcus* phages to test design principles of phage cocktails and determine the phenotypic effects of evolving resistance. We show that cocktails of two or three unrelated phages often, but not always, prevented the growth of phage-resistant mutants. Cocktails of related phages were generally not effective at preventing the growth of phage-resistant *Enterococcus*. Many of the mutations that provide *Enterococcus* resistance to phage infection involved exopolysaccharide synthesis genes for all types of phages tested. Further, we show that evolving resistance to phage infection can alter *Enterococcus* susceptibility to vancomycin. This work will help to inform phage cocktail design for future phage therapy applications.

INTRODUCTION

Antibiotic resistant bacterial infections have emerged as a major health crisis. Overuse of antibiotics has led to rising rates of antibiotic resistance, therefore steps must be taken to find alternative therapies. The ESKAPE pathogens are a list of six high-priority antibiotic resistant bacteria, and among them is *Enterococcus*, a gram-positive bacterium responsible for many hospital-acquired infections (Pendleton, Gorman, and Gilmore 2013).

Rates of vancomycin resistant *Enterococcus* (VRE) infections have been rising; one in three hospital acquired *Enterococcus* infections is vancomycin resistant (Agudelo Higuita and Huycke 2014b). Bacteriophage (phage) therapy is an alternative treatment to antibiotics that has shown promise *in vitro* and in animal models for treating *Enterococcus* (Yoong et al. 2004b; Chatterjee et al. 2019; Khalifa et al. 2015a). However, phage therapy is rarely used and only as a last resort because basic research into the safety, mechanisms, and best practices is lacking.

Although phages are posited as a solution for antibiotic resistance, bacteria can also evolve resistance to phage infection. Bacteria exist in a constant evolutionary battle with phage, and thus have evolved many systems to resist phage infection, including preventing phage binding, restriction modification systems, CRISPR-Cas9, an abortive infection (Dy et al. 2014). Given strong selective pressure from a single phage, bacteria often quickly evolve resistance to phage in laboratory settings (Chao, Levin, and Stewart 1977). Phage evolve to combat and circumvent these resistance mechanisms, which might provide phage therapy two advantages over antibiotics (Samson et al. 2013b). One, during phage therapy, if bacteria evolve resistance, phage can also evolve to overcome that resistance. However, it remains unclear whether there is time or opportunity to evolve during phage therapy in humans. Two, this evolutionary battle is ongoing in nature, so isolating phages from the environment should provide an endless reservoir of phages with mutations that allow them to circumvent resistance mechanisms. But this endless diversity also creates the challenge of understanding which phages and which genotypes will be most effective for phage therapy.

Additionally, the phenotypic effects of evolving resistance to phages is poorly understood. For one, evolving resistance to a single phage could result in cross-resistance to other phages (R. C. T. Wright et al. 2018). The networks of hosts that a phage will infect have been seen to be nested and structured around host receptor, but patterns of cross resistance have not been well characterized (Flores et al. 2011b; Lennon et al. 2007). Another effect that phage-resistance might have is altering susceptibility to antibiotics (Chan et al. 2016). Vancomycin resistant *Enterococcus* that evolved resistance to phage was seen to have greatly increased susceptibility to vancomycin, in addition to mouse gut colonization defects (Chatterjee et al. 2019). Increasing antibiotic susceptibility would be an ideal outcome of evolving resistance to phage, but more work needs to be done to determine how often these fitness tradeoffs occur.

Phage therapy is often administered as a cocktail of phages. This is done because phages often have narrow host ranges, so multiple phages are used to ensure all of the target bacteria will be killed (Flores et al. 2011b). Theoretically, using a cocktail of phages could also decrease the chance that a phage-resistant mutant can arise, since multiple mechanisms of resistance would need to evolve simultaneously. Similarly, combinations of antibiotics are used to treat tuberculosis infections, and combinations of antivirals are used to treat HIV (Günthard et al. 2016; Tornheim and Dooley 2019). However, the efficacy of phage cocktails has not been thoroughly explored, and instead it is taken for granted that a cocktail of phages will be more effective than a single phage. Thus, there are no design principles for crafting effective phage cocktails. Logic might suggest that a diverse collection of phages should be used to kill a broader range of hosts and to prevent

resistance arising to all phages by a common mechanism. However, this logic still needs to be tested in a wide range of systems.

Phage therapy has not been used to treat *Enterococcus* infections of humans specifically, but it has shown promise in *in vitro* and *in vivo* mouse experiments (Khalifa et al. 2015a, 2016; Chatterjee et al. 2019). Additionally, *Enterococcus* phages have been shown to be effective at disrupting *Enterococcus* in biofilms, which are generally much harder to treat than planktonic cells because antibiotics have trouble penetrating biofilms (Khalifa et al. 2015a). *Enterococcus* phages have also been used to treat humans. Two phage cocktails sold by the Eliava Institute of Bacteriophage, Microbiology, and Virology in Georgia were shown to contain abundant *Enterococcus* phages (Villarroel et al. 2017). *Enterococcus* phages have been evaluated in cocktails previously, and were shown to be more effective than single phages at preventing the growth of resistant *Enterococcus* mutants (Khalifa et al. 2018). While encouraging, this is only one example of cocktail design and does not offer any insight into cocktail design principles.

We utilized our collection of fourteen *Enterococcus* phages to test the efficacy of different combinations of phages at bacterial killing and preventing the growth of phage-resistant mutants. In doing so, we isolated and sequenced phage-resistant *Enterococcus* mutants that grew in the presence of single phages and phage cocktails. To determine the extent of cross resistance that occurs when evolving resistance to *Enterococcus* phages, phage-resistant *Enterococcus* mutants had their susceptibility to all other phages measured. Finally, we measured the vancomycin susceptibility of phage-resistant mutants to see if evolving resistance to phage affected antibiotic susceptibility. Overall these experiments will inform future *Enterococcus* phage cocktail design.

RESULTS

Phage cocktails prevent arising of resistant mutants

The strain *Enterococcus faecalis* Yi-6 was chosen for phage cocktail because it was susceptible to all 15 phages in our collection. When *E. faecium* Yi6 was infected with a single phage at a high MOI (**Table S5.1**) in liquid media, phage-resistant bacteria usually grew back within 72 hours. (**Figure 5.1**). Related *Siphoviridae* phages vB_OCPT_SDS1 and vB_OCPT_SDS2 were the most effective single phages for preventing the growth of phage-resistant mutants. When *E. faecium* Yi6 was infected with combinations of two phages, some combinations consistently prevented the growth of phage-resistant mutants, while other combinations consistently failed to do so. In particular, combinations of phages involving the *Brockvirinae* phage vB_OCPT_Ben along with a *Siphoviridae* or a *Podoviridae* phage always prevented the growth of a resistant mutant. All combinations that prevented the growth of phage-resistant mutants were unrelated phages, however not all combinations of unrelated phages were successful. Combinations of genetically related phages always allowed the growth of phage-resistant mutants. Combinations of three phages were able to prevent the growth of phage-resistant mutants in 4/5 combinations that included at least two unrelated phages. However, adding a third phage did not improve on combinations that failed as two phages. All successful three-phage combinations contained two phages that worked together as two-phage cocktails.

Enterococcus resist diverse phages by exopolysaccharide mutations

Phage resistant *E. faecalis* Yi6 and *E. faecalis* DP11 mutants were generated against a diverse collection of phages. Mutations in the Epa exopolysaccharide synthesis loci were the most common mutations observed (5/6 phage resistant mutants) (**Table 5.1**). Most of

these phage-resistant mutants had a single point mutation or nonsense mutation in a single gene in the Epa locus and did not contain any other mutations. Phage-resistant mutants were isolated by the ability to grow in the presence of an infecting phage, but in some cases, incomplete resistance was observed. In these cases, mutations provided enough resistance for *Enterococcus* to be isolated in our assay, but the infecting phage still shows signs of lysis upon a subsequent drop assay.

After generating *Enterococcus* mutants that were resistant to a single phage, these mutants were tested for susceptibility against other phages to test for cross resistance (**Figure 5.2**). Overall, phage-resistant *Enterococcus* mutants generated against a single phage often were cross-resistant to several other phages. On average, *Enterococcus* mutants that evolved resistance to a single phage were susceptible to 60 % fewer phages than the wild type strain. Cross resistance was observed against closely related phages and for unrelated phages. For example, when *E. faecalis* DP11 evolved resistance to a Siphoviridae phage, *vB_OCPT_SDS1*, it was also resistant to seven out of eight *Herelleviridae* phages that were able to infect the wild type.

Resistance to phage can modulate antibiotic susceptibility

We tested to see if *Enterococcus* phage-resistant mutants have altered susceptibility to vancomycin. *E. faecalis* strains DP11 and Yi6 were initially susceptible to vancomycin while E. faecalis V587 was vancomycin resistant. Phage-resistant *Enterococcus* mutants often had similar levels of susceptibility, but in a few cases, resistance was seen to increase or decrease (**Figure 5.3**). One of the E. faecalis V587 phage resistant mutants was more susceptible than the wild type to vancomycin at 100 ug/ml. The phage-resistant mutants generated from the vancomycin sensitive strains, *E. faecalis* Yi6 and DP11 either had no

change in vancomycin sensitivity or were more resistant. Many phage-resistant *Enterococcus* mutants did not grow up to the same OD600 as the wild type, indicating that these mutants have general growth defects.

DISCUSSION

Phage cocktails have been shown to be effective at killing bacteria and preventing the growth of resistant mutants, but little focus has been placed on the design principles for effective cocktails. Here, we show that phage cocktail design is an important consideration because some combinations of phages consistently prevent the growth of phage-resistant mutants while others do not.

While using phage cocktails is a standard practice in phage therapy, there is no standard for how many phages should be used. The pyophage (PYO) cocktail from the Georgian Eliava Institute of Bacteriophage has been shown to contain approximately thirty different phages, however these cocktails are designed to target multiple different bacterial hosts (Villarroel et al. 2017). Recent uses of phage therapy designed to target a single bacteria generally include between one and six phages(A. Wright et al. 2009; Schooley et al. 2017; Duplessis et al. 2017). Often there is no reasoning behind the number of phages chosen to administer during phage therapy. Here, we show that combinations of two phages are enough to prevent the growth of phage-resistant mutants. Using more than two diverse phages in a cocktail would increase the chances of choosing two phages that displayed synergy in preventing the growth of phage-resistant mutants.

Several approaches exist for optimizing phage cocktail design. Experimental evolution of a phage can result in mutant phages with expanded host ranges, which can be utilized in phage cocktails (D. Kelly et al. 2011; Burrowes et al. 2019). Another approach is

to use phage-resistant hosts to isolate new phages with complementary host ranges (Gu et al. 2012). Synthetic approaches can also be effective, such as using site-directed mutants to discover phages that bind to different sites which could complement each other in a phage cocktail (Filippov et al. 2011). Our experiments relied on the diversity of our phage collection to create cocktails, but many bacterial hosts lack large catalogs of characterized phages. Evolutionary or synthetic approaches would be useful for expanding phage catalogs for specific hosts.

Host range is an important consideration for choosing phages to use in a phage cocktail. Broad host ranges phages can infect multiple genera, while a narrow host range phage is seemingly specific to a single strain (Flores et al. 2011a). The full scope of host range is difficult to determine due to limitations in the number of bacterial strains that need to be isolated. Phage infection networks have been seen to be highly nested and structured around the host receptor they utilize (Flores et al. 2011b; R. C. T. Wright et al. 2018). However, previous work in *Synechococcus* showed no correlation between cross-resistance patterns and genetic similarity between phages (Stoddard, Martiny, and Marston 2007). Since the receptors of the phages used in this study are largely unknown, one explanation for the high degree of cross-resistance is that they utilize a common receptor. Evidence exists that the *Brockvirinae* phages utilized in this study bind to cell wall exopolysaccharides (Chatterjee et al. 2019; Wandro et al. 2019). If phage infection for multiple types of *Enterococcus* phages is mediated by binding to exopolysaccharides, then broad patterns of cross-resistance, such as we observed, might be expected to occur.

Phages that display synergy with antibiotics would be ideal for phage therapy. Synergy can occur if the evolution of resistance to phage comes at the cost of increased

susceptibility to antibiotics, or vice versa. The most famous example of this is a *Pseudomonas aeruginosa* phage that that binds to an antibiotic efflux pump, resulting in a fitness trade off where bacteria can evolve resistance to phage by losing the antibiotic efflux pump (Chan et al. 2016). *Enterococcus* have previously been seen to evolve resistance to phage infection through mutations in the Epa locus, which has been seen to increase susceptibility to vancomycin in previously vancomycin resistant strains (Chatterjee et al. 2019; Wandro et al. 2019; Singh, Lewis, and Murray 2009; Ho et al. 2018). The mechanism by which this fitness trade-off occurs is not clear; vancomycin targets peptidoglycan and the Epa locus regulates exopolysaccharide synthesis, so both interactions occur at the cell wall (Thurlow, Thomas, and Hancock 2009; Hancock and Gilmore 2002; Fang Teng et al. 2009). In this study, evolution of phage resistance in *Enterococcus* was seen to have minimal effects on vancomycin sensitivity, even though many of the mutations also occurred in the Epa locus. The mechanism by which Epa mutations provide resistance to phage infection is unclear, and different mutations may have differential effects on antibiotic susceptibility. In addition, the genetic background of the *Enterococcus* strain likely affects the phenotypic outcomes of the evolution of phage resistance. In some cases, evolving resistance to phage infection may alter antibiotic susceptibility and other colonization phenotypes. Synergistic outcomes of phage-host co-evolution that result in better treatment options are far from guaranteed, but the enormous numbers of understudied phage-host interactions leave room the discovery of phages that co-evolve with their pathogenic bacterial hosts in ways that make infection treatment more tractable.

MATERIALS AND METHODS

Mutant Generation in Liquid Cultures

Enterococcus colonies were grown overnight with 1xBHI broth in a shaker at 37 °C. The next day, all cultures were diluted to .05 OD600 in fresh BHI, where 190 microliters of *Enterococcus* and 10 microliters of a highly concentrated individual bacteriophage stock were added into each well. 96-well plates filled with bacteria-bacteriophage pairs were then incubated at 37C inside a spectrophotometer, where their OD600 was taken every 10 minutes for 24-48 hours. During the 1-2-day incubation, any Enterococcus sp. culture which was initially annihilated to the background optical density of only the media and well by themselves, but subsequently saw a detectable rebound of growth, were streaked onto 1.5% Agar plates infused with BHI. Colonies which grew upon these streak plates were considered bacteriophage resistant mutants of the bacteriophage they were inoculated with and were used in subsequent mutant assays.

Mutant Generation on Agar Plates

Enterococcus colonies were grown overnight with 1xBHI broth in a shaker at 37 °C. The next day, all cultures were diluted to .05 OD600, in 2.5mL of .35% Low Melting-Point Agarose infused with BHI and inoculated with 10 microliters of phage stock. The ~2.5mL's of Enterococcus-bacteriophage mixture was then poured on top of a 1.5% Agarose plate infused with BHI and allowed to solidify. After 24-48 hours, visible Enterococcus colonies displaying bacteriophage resistance by growing on top of the agar were picked and re-streaked for subsequent mutant assays.

Bacteriophage Host Infectivity Range

The susceptibility of wild-type and mutant *Enterococcus* isolates were tested using a spot assay on agar plates. Colonies picked from streak plates were cultured overnight with BHI broth in a shaker at 37C. The next day, *Enterococcus* cultures were diluted to .05 OD600 in ~2.5mL of .35% Low Melting-Point Agarose infused with BHI and poured onto a 1.5% Agar plate infused with BHI. After waiting ~30 minutes for .35% Low Melting-Point Agarose infused with BHI to solidify, each of our 15 *Enterococcus* phages were added in 6 microliter spots on top of the solidified .35% Low Melting-Point Agarose, and plates were incubated at 37C. The negative control received 6 microliters of SM buffer which the phages are stored in. After 24 hour of incubation plates were examined and plaque-like clearings which appeared at the site of 6 microliter phage droplet inoculation were recorded.

Sequencing mutants

DNA was extracted from Enterococcus mutants using the Quick-DNA Microprep Kit (Zymo #D3020). Before Enterococcus DNA extraction, lysozyme was added to lysis buffer at a concentration of 100ug/ml and incubated at 37 °C for 30 minutes. Libraries were prepared using a scaled-down protocol with the Illumina Nextera enzyme (Baym et al., 2015). 75 bp short read length paired-end sequencing was performed on the Illumina NextSeq using the Mid Output v2 reagents. Approximately 1 million reads were obtained per sample, giving about 10-fold coverage across the *Enterococcus* genome.

Vancomycin Susceptibility

To determine whether becoming a bacteriophage resistant mutant affected Vancomycin susceptibility or resistance, we cultured wild-type and mutant *Enterococcus* across a spectrum of relevant vancomycin concentrations. Wild-type and Mutant

Enterococcus colonies were grown overnight with BHI broth in a shaker at 37 °C. The next day, all cultures were diluted to .05 OD600 in fresh BHI. In a 96 well plate, 190 microliters of *Enterococcus* and 10 microliters of Vancomycin in a gradient from .2 to 200 micrograms/mL was added in triplicate to wells and incubated in a spectrophotometer at 37 °C for 10 hours.

Phage cocktails

Cocktails consisting of one, two, and three combinations were tested against *E. faecalis* Yi6. A single colony was grown overnight with BHI broth in a shaker at 37 °C. The next day, bacterial cultures were diluted to .05 OD600 in fresh BHI, where 190 microliters of *Enterococcus* and 10 microliters of phage were combined in a single well inside a 96 well plate. Phages were always added at the highest concentration available (**Table S5.1**). For two-phage cocktails, 5 microliters of two unique phage stocks were added. To conduct three-phage cocktails, 3.33 microliters of 3 unique phage stocks were added. Plates were incubated for 72 hours at 37 °C inside of a spectrophotometer. At 48 hours an additional 100 microliters of fresh BHI media was added to each well to avoid desiccation.

FIGURES AND TABLES

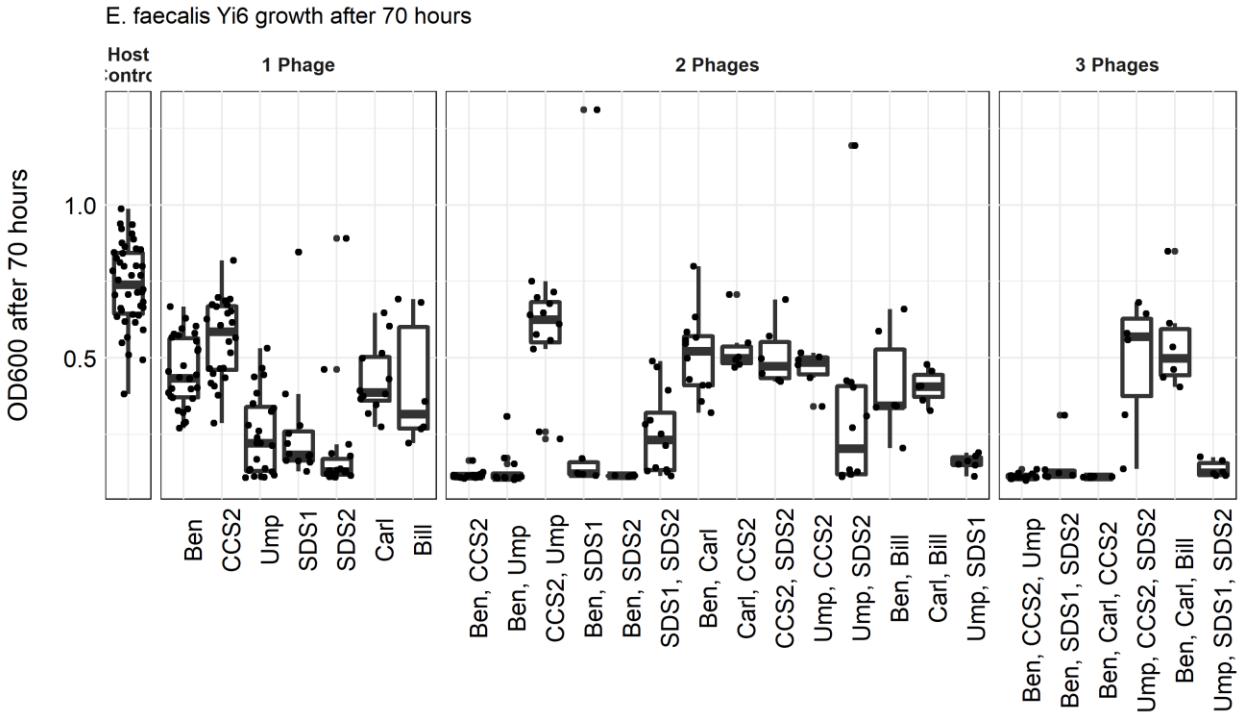


Figure 5.1. Phage cocktails prevent growth of resistant mutants. *E. faecalis* Yi6 growth in liquid culture with one, two, or three phages after 70 hours. Growth at 70 hours indicates a resistant mutant was able to evolve and grow.

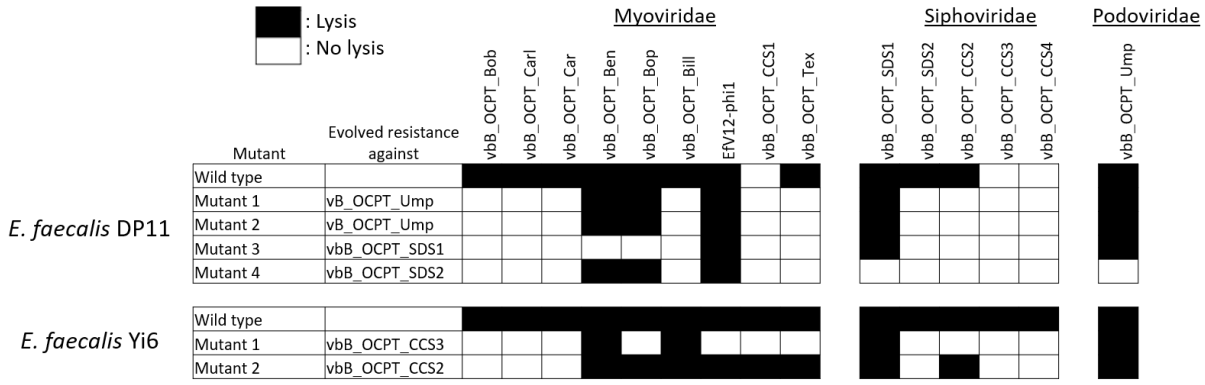


Figure 5.2. Cross-resistance patterns. Phage susceptibility of *Enterococcus* that evolved resistance to a single phage was tested against the collection of *Enterococcus* phages by drop assay.

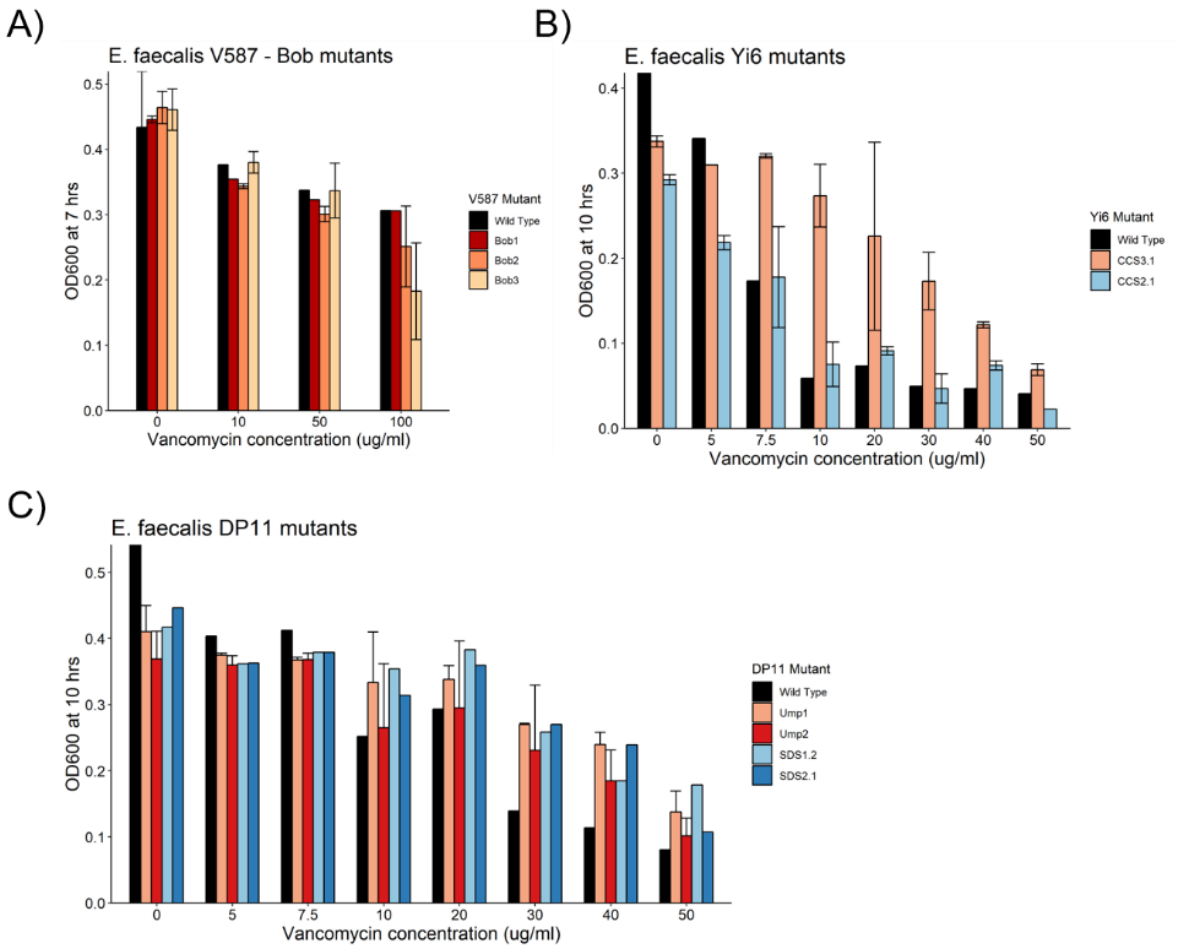


Figure 5.3. Vancomycin susceptibility of *Enterococcus* mutants. Phage-resistant mutants generated for **A)** vancomycin resistant strain E. faecalis V587 **B)** E. faecalis Yi6 and **C)** E. faecalis DP11.

Table 5.1. Phage information

Name	Family	Genus	Genome size
vB_OCPT_Bob	Brockvirinae	Kochiodavirus	150k
vB_OCPT_Car	Brockvirinae	Kochiodavirus	150k
vB_OCPT_Carl	Brockvirinae	Kochiodavirus	150k
EfV12-phi1	Brockvirinae	Wandervirus	150k
vB_OCPT_Ben	Brockvirinae	Wandervirus	150k
vB_OCPT_Bop	Brockvirinae	Wandervirus	150k
vB_OCPT_Bill	Brockvirinae	Wandervirus	150k
vB_OCPT_CCS1	Brockvirinae	Wandervirus	150k
vB_OCPT_Tex	Brockvirinae	Wandervirus	150k
vB_OCPT_SDS1	Siphoviridae	vB_EfaS_IME198	57k
vB_OCPT_SDS2	Siphoviridae	vB_EfaS_IME198	57k
vB_OCPT_CCS2	Siphoviridae	vB_EfaS_IME198	57k
vB_OCPT_CCS3	Siphoviridae	vB_EfaS_IME198	57k
vB_OCPT_Toy	Siphoviridae	vB_EfaS_IME198	57k
vB_OCPT_CCS4	Siphoviridae	vB_EfaS_AL3	40k
vB_OCPT_Ump	Podoviridae	vB_EfaP_IME199	18k

Table 5.2. Mutations in Enterococcus providing phage resistance. Genes that are part of the Epa exopolysaccharide synthesis locus are denoted with (Epa).

Host	Mutant	Phage	Mutated gene(s)	Mutation
***E. faecalis* DP11**	Mutant 1	vB_OCPT_Ump	(Epa) WgbU UDP-N-acetylglucosamine 4-epimerase	G11V (60%)
***E. faecalis* DP11**	Mutant 2	vB_OCPT_Ump	(Epa) WecA	E249*
***E. faecalis* DP11**	Mutant 3	vB_OCPT_SDS1	(Epa) WgbU UDP-N-acetylglucosamine 4-epimerase	G279E
***E. faecalis* DP11**	Mutant 4	vB_OCPT_SDS2	(Epa) TagF gene. glycerol glycerophosphotransferase	A113E
***E. faecalis* Yi6**	Mutant 1	vB_OCPT_CCS3	(Epa) exopolysaccharide biosynthesis; aspartate aminotransferase	W7* ; P365S
***E. faecalis* Yi6**	Mutant 2	VB_OCPT_CCS2	cell division ABC transporter permease FtsX	P118A

Table S5.1. Phage titers. All phages tittered on *E. faecalis* Yi6.

Phage	Titer (PFU/ml)
vB_OCPT_Ben	4x10^10
vB_OCPT_Carl	4x10^8
vB_OCPT_Bill	2x10^8
vB_OCPT_SDS1	1.8x10^10
vB_OCPT_SDS2	2x10^8
vB_OCPT_CCS2	3x10^9
vB_OCPT_Ump	7.7x10^8

SUMMARY AND FUTURE DIRECTIONS

Phages have the potential to revolutionize the way we interact with microbial communities. In the years since this thesis work began, rapid progress has been made toward our understanding of the human microbiome and the phages within it. This is due in part to improvements both in sequencing technologies and bioinformatics methods (Levy and Myers 2016; Kono and Arakawa 2019; Roumpeka et al. 2017). Increasingly, metagenomics is being used to characterize microbial communities, which has allowed the study of non-bacterial members of the microbiome. However, many questions about phages in the microbiome remain unanswered: How do phages influence the composition of bacteria in the microbiome? Can phages be used to intentionally alter the composition of the microbiome? What are the best practices for using phages as therapeutics? If we can answer these questions, we will have a much more complete understanding of the human microbiome and also have powerful tools for manipulating it. This dissertation begins to answer these questions by focusing on human-associated *Enterococcus* and how it interacts with its phages.

Studying the development of the microbiome early in life proved to be a valuable model for antibiotic-driven Enterococcal blooms. The microbiome of preterm infants was shown to be highly personalized, yet each were dominated by fast-growing, facultative anaerobes due to antibiotic exposure (Wandro, Osborne, et al. 2017). Since the source of *Enterococcus* infections often can be traced to strains colonizing the gut, it is important to know when and how these overgrowths occur (Ubeda et al. 2010b). Infants have a distinct microbiome from adults that is less robust to antibiotic perturbation (Sharon et al. 2013; Gibson et al. 2016). Since antibiotics are often a necessity in infants, future work will need

to focus on ways to mitigate the negative effects of antibiotics on the infant microbiome. Of particular concern is the permanent loss of beneficial bacterial strains, such as *Bifidobacteria infantis* (Karav et al. 2016; Dethlefsen and Relman 2011b). However, there is evidence that supplementation of specific *B. infantis* strains can be effective (Frese et al. 2017). Supplementation with beneficial strains, especially after antibiotic usage, could prevent the overgrowth of opportunistic pathogens such as *Enterococcus*.

The use of metabolomics to study the microbiome has continued to unveil new insights into the interactions between humans and microbes. In this dissertation, untargeted metabolomics showed close associations with the overall composition of bacteria in the microbiome, but no universal biomarkers of disease or dysbiosis were observed. The microbiome and metabolome are both highly personalized, thus finding shared signals across patients in a noisy background such as fecal samples remains difficult. Another challenge for untargeted metabolomics is the lack of annotations for most metabolites (Viant et al. 2017). Progress is ongoing toward improving metabolite annotations and networking related compounds (Wang et al. 2016). Future work will hopefully move towards standardizing the use of metabolomics in microbiome studies, so that signals can be compared among datasets.

Understanding how bacteria and phage co-evolve is an important consideration for phage therapy applications. When bacteria evolve resistance to phage, there can be fitness trade-offs that decrease bacterial virulence (Chan et al. 2016; Chatterjee et al. 2019). However, this outcome has only been observed in a few cases and is likely to be specific to a minority of phage-bacteria interactions. Further, environmental conditions have shown to alter the evolutionary strategies of bacterial resistance to phage (Alseth et al. 2019).

Experimental co-evolution as used in this dissertation can be a powerful tool for characterizing bacteriophage interactions and for finding phages that produce these desirable outcomes for phage therapy. However, these experiments are not feasible on the timescale of clinical applications, so learning to predict the outcomes of co-evolution between bacteria and phages would greatly improve the potential of phages as therapeutics. Future work characterizing coevolution between bacteria and phage in diverse systems and diverse environments will be required before we are able to predict the outcomes.

Phage therapy has made great strides in in recent years (Kortright et al. 2019b). Programs such as the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) have expanded phage banks with large-scale isolation and characterization of new phages – some of them have already been engineered and implemented in treatments as cocktails (Hanauer et al. 2017; Dedrick et al. 2019). Phage therapy is beginning to be applied clinically in select compassionate-use-exemption cases in the United States and Western Europe (Kortright et al. 2019a). Several studies have laid the groundwork for using *Enterococcus* phages therapeutically (Khalifa et al. 2016, 2015a, 2018; Chatterjee et al. 2019). We improve on this groundwork by characterizing multiple new *Enterococcus* phages, showing how they co-evolve with their hosts, and demonstrating how they should be combined in phage cocktails. While we utilized a large collection of *Enterococcus* relative to most studies, we have likely not even scratched the surface of *Enterococcus* phage diversity that exists in nature. Future work will be required to isolate and characterize new *Enterococcus* phages, to see if the trends we observe are

generalizable beyond our collection of phages. Phage therapy is still in early stages but could eventually be part of the solution to the problem of antibiotic resistance.

REFERENCES

Agudelo Higuita, Nelson I., and Mark M Huycke. 2014a. *Enterococcal Disease, Epidemiology, and Implications for Treatment*. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. http://www.ncbi.nlm.nih.gov/pubmed/24649504.

Alseth, Ellinor O, Elizabeth Pursey, Adela M Luján, Isobel McLeod, Clare Rollie, and Edze R Westra. 2019. "Bacterial Biodiversity Drives the Evolution of CRISPR-Based Phage Resistance in Pseudomonas Aeruginosa." *BioRxiv*, March. Cold Spring Harbor Laboratory, 586115. https://doi.org/10.1101/586115.

Andersen, Heidi, Natalia Connolly, Hansraj Bangar, Mary Staat, Joel Mortensen, Barbara Deburger, and David B Haslam. 2016. "Use of Shotgun Metagenome Sequencing To Detect Fecal Colonization with Multidrug-Resistant Bacteria in Children." *Journal of Clinical Microbiology* 54 (7). American Society for Microbiology: 1804–13. https://doi.org/10.1128/JCM.02638-15.

Asnicar, Francesco, Serena Manara, Moreno Zolfo, Duy Tin Truong, Matthias Scholz, Federica Armanini, Pamela Ferretti, et al. 2017. "Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling." Edited by Jack A. Gilbert. *MSystems* 2 (1). https://doi.org/10.1128/mSystems.00164-16.

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, et al. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5). Elsevier: 690–703. https://doi.org/10.1016/j.chom.2015.04.004.

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 (5): 455–77. https://doi.org/10.1089/cmb.2012.0021.

Barylski, Jakub, Francois Enault, Bas E Dutilh, Margo BP Schuller, Robert A Edwards, Annika Gillis, Jochen Klumpp, et al. 2018. "Analysis of Spounaviruses as a Case Study for the Overdue Reclassification of Tailed Bacteriophages." *BioRxiv*, February. Cold Spring Harbor Laboratory, 220434. https://doi.org/10.1101/220434.

Bastin, David A, Gordon Stevenson, Peter K Brown, Antje Haase, and Peter R Reeves. 1993. "Repeat Unit Polysaccharides of Bacteria: A Model for Polymerization Resembling That of Ribosomes and Fatty Acid Synthetase, with a Novel Mechanism for Determining Chain Length." *Molecular Microbiology* 7 (5): 725–34. https://doi.org/10.1111/j.1365-2958.1993.tb01163.x.

Bäumler, Andreas J, and Vanessa Sperandio. 2016. "Interactions between the Microbiota and Pathogenic Bacteria in the Gut." *Nature* 535 (7610). NIH Public Access: 85–93. https://doi.org/10.1038/nature18849.

Baym, Michael, Sergey Kryazhimskiy, Tami D Lieberman, Hattie Chung, Michael M Desai, and Roy Kishony Kishony. 2015. "Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes." *PLoS ONE* 10 (5): 1–15. https://doi.org/10.1371/journal.pone.0128036.

Bentley, Stephen D., David M. Aanensen, Angeliki Mavroidi, David Saunders, Ester Rabbinowitsch, Matthew Collins, Kathy Donohoe, et al. 2006. "Genetic Analysis of the Capsular Biosynthetic Locus from All 90 Pneumococcal Serotypes." *PLoS Genetics* 2 (3).

Public Library of Science: e31. https://doi.org/10.1371/journal.pgen.0020031.
Bohannan, B.J.M., and R.E. Lenski. 2000. "Linking Genetic Change to Community Evolution: Insights from Studies of Bacteria and Bacteriophage." *Ecology Letters* 3 (4). Blackwell Science Ltd: 362–77. https://doi.org/10.1046/j.1461-0248.2000.00161.x.
Bokulich, Nicholas A, Jennifer Chung, Thomas Battaglia, Nora Henderson, Melanie Jay, Huilin Li, Arnon D Lieber, et al. 2016. "Antibiotics, Birth Mode, and Diet Shape Microbiome Maturation during Early Life." *Science Translational Medicine* 8 (343). American Association for the Advancement of Science: 343ra82. https://doi.org/10.1126/scitranslmed.aad7121.
Bolduc, Benjamin, Ho Bin Jang, Guilhem Doulcier, Zhi-Qiang You, Simon Roux, and Matthew B. Sullivan. 2017. "VConTACT: An iVirus Tool to Classify Double-Stranded DNA Viruses That Infect *Archaea* and *Bacteria*." *PeerJ* 5 (May). PeerJ Inc.: e3243. https://doi.org/10.7717/peerj.3243.
Bos, Lieuwe D J, Simone Meinardi, Donald Blake, and Katrine Whiteson. 2016. "Bacteria in the Airways of Patients with Cystic Fibrosis Are Genetically Capable of Producing VOCs in Breath." *Journal of Breath Research* 10 (4): 047103 %U http://stacks.iop.org/1752-7163/10/i=4/a.
Bos, Lieuwe D J, Peter J Sterk, and Marcus J Schultz. 2013. "Volatile Metabolites of Pathogens: A Systematic Review." *PLoS Pathog* 9 (5): e1003311 %U http://dx.doi.org/10.1371/journal.ppat.
Bos, Lieuwe D J, Hans Weda, Yuanyue Wang, Hugo H Knobel, Tamara M E Nijsen, Teunis J Vink, Aeilko H Zwinderman, Peter J Sterk, and Marcus J Schultz. 2014. "Exhaled Breath Metabolomics as a Noninvasive Diagnostic Tool for Acute Respiratory Distress Syndrome." *The European Respiratory Journal* 44 (1): 188–97.
Bouvier, T, and P A Giorgio. 2007. "Key Role of Selective Viral-Induced Mortality in Determining Marine Bacterial Community Composition" 9: 287–97. https://doi.org/10.1111/j.1462-2920.2006.01137.x.
Bradford, Patricia A. 2018. "Epidemiology of Bacterial Resistance." In *Antimicrobial Resistance in the 21st Century*, 299–339. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-78538-7_10.
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 (5): 367–73. https://doi.org/10.1016/j.resmic.2008.04.006.
Brooks, Brandon, Matthew R. Olm, Brian A. Firek, Robyn Baker, Brian C. Thomas, Michael J. Morowitz, and Jillian F. Banfield. 2017. "Strain-Resolved Analysis of Hospital Rooms and Infants Reveals Overlap between the Human and Room Microbiome." *Nature Communications* 8 (1). Nature Publishing Group: 1814. https://doi.org/10.1038/s41467-017-02018-w.
Brüssow, Harald. 2016. "How Stable Is the Human Gut Microbiota? And Why This Question Matters." *Environmental Microbiology* 18 (9): 2779–83. https://doi.org/10.1111/1462-2920.13473.
Buffie, Charlie G., Irene Jarchum, Michele Equinda, Lauren Lipuma, Asia Gobourne, Agnes Viale, Carles Ubeda, Joao Xavier, and Eric G. Pamer. 2012. "Profound Alterations of Intestinal Microbiota Following a Single Dose of Clindamycin Results in Sustained Susceptibility to Clostridium Difficile-Induced Colitis." Edited by B. A. McCormick.

Infection and Immunity 80 (1): 62–73. https://doi.org/10.1128/IAI.05496-11.
Buffington, Shelly A., Gonzalo Viana Di Prisco, Thomas A. Auchtung, Nadim J. Ajami, Joseph F. Petrosino, and Mauro Costa-Mattioli. 2016. "Microbial Reconstitution Reverses Maternal Diet-Induced Social and Synaptic Deficits in Offspring." *Cell* 165 (7): 1762–75. https://doi.org/10.1016/j.cell.2016.06.001.
Burrowes, Ben, Ian Molineux, Joe Fralick, Ben H. Burrowes, Ian J. Molineux, and Joe A. Fralick. 2019. "Directed in Vitro Evolution of Therapeutic Bacteriophages: The Appelmans Protocol." *Viruses* 11 (3). Multidisciplinary Digital Publishing Institute: 241. https://doi.org/10.3390/v11030241.
Cajka, Tomas, and Oliver Fiehn. 2016. "Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics." *Analytical Chemistry* 88 (1). American Chemical Society: 524–45. https://doi.org/10.1021/acs.analchem.5b04491.
Canchaya, Carlos, Ghislain Fournous, Sandra Chibani-Chennoufi, Marie Lise Dillmann, and Harald Brüssow. 2003. "Phage as Agents of Lateral Gene Transfer." *Current Opinion in Microbiology* 6 (4): 417–24. http://www.ncbi.nlm.nih.gov/pubmed/12941415.
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 (5). BioMed Central: 335–36. https://doi.org/10.1038/nmeth.f.303.
Carmody, Lisa A, Jiangchao Zhao, Linda M Kalikin, William LeBar, Richard H Simon, Arvind Venkataraman, Thomas M Schmidt, Zaid Abdo, Patrick D Schloss, and John J LiPuma. 2015. "The Daily Dynamics of Cystic Fibrosis Airway Microbiota during Clinical Stability and at Exacerbation." *Microbiome* 3: 12 %U http://dx.doi.org/10.1186/s40168-015-0074-9.
Cassir, Nadim, Samia Benamar, Jacques Bou Khalil, Olivier Croce, Marie Saint-Faust, Aurélien Jacquot, Matthieu Million, et al. n.d. "Clostridium Butyricum Strains and Dysbiosis Linked to Necrotizing Enterocolitis in Preterm Neonates" 61 (7): 1107–15. Accessed November 3, 2015. https://doi.org/10.1093/cid/civ468.
Chan, Benjamin K, Mark Sistrom, John E Wertz, Kaitlyn E Kortright, Deepak Narayan, and Paul E Turner. 2016. "Phage Selection Restores Antibiotic Sensitivity in MDR Pseudomonas Aeruginosa." *Scientific Reports* 6 (May). Nature Publishing Group: 26717. https://doi.org/10.1038/srep26717.
Chan, Benjamin K, Paul E Turner, Samuel Kim, Hamid R Mojibian, John A Elefteriades, and Deepak Narayan. 2018. "Phage Treatment of an Aortic Graft Infected with Pseudomonas Aeruginosa." *Evolution, Medicine, and Public Health* 2018 (1). Oxford University Press: 60–66. https://doi.org/10.1093/emph/eoy005.
Chang, Pamela V, Liming Hao, Stefan Offermanns, and Ruslan Medzhitov. 2014. "The Microbial Metabolite Butyrate Regulates Intestinal Macrophage Function via Histone Deacetylase Inhibition." *Proceedings of the National Academy of Sciences* 111 (6): 2247–52. https://doi.org/10.1073/pnas.1322269111.
Chao, Lin, Bruce R. Levin, and Frank M. Stewart. 1977. "A Complex Community in a Simple Habitat: An Experimental Study with Bacteria and Phage." *Ecology* 58 (2). Ecological Society of America: 369–78. https://doi.org/10.2307/1935611.
Charbonneau, Mark R., Laura V. Blanton, Daniel B. Digiulio, David A. Relman, Carlito B. Lebrilla, David A. Mills, and Jeffrey I. Gordon. 2016. "A Microbial Perspective of Human

Developmental Biology." *Nature*. https://doi.org/10.1038/nature18845.
Chatterjee, Anushila, Cydney N Johnson, Phat Luong, Karthik Hullahalli, Sara W McBride, Alyxandria M Schubert, Kelli L Palmer, Paul E Carlson, and Breck A Duerkop. 2019. "Bacteriophage Resistance Alters Antibiotic Mediated Intestinal Expansion of Enterococci." *BioRxiv*, January. Cold Spring Harbor Laboratory, 531442. https://doi.org/10.1101/531442.
Chevreux, B, T Wetter, and S Suhai. 1999. "Genome Sequence Assembly Using Trace Signals and Additional Sequence Information." *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) '99, GCB, Hannover, Germany.*, no. 1995: 45–56. https://doi.org/10.1.1.23/7465.
Clark, Reese H, Barry T Bloom, Alan R Spitzer, and Dale R Gerstmann. 2006. "Reported Medication Use in the Neonatal Intensive Care Unit: Data From a Large National Data Set." *Pediatrics* 117 (6): 1979–87. https://doi.org/10.1542/peds.2005-1707.
Cole, James R, Qiong Wang, Jordan A Fish, Benli Chai, Donna M McGarrell, Yanni Sun, C Titus Brown, Andrea Porras-Alfaro, Cheryl R Kuske, and James M Tiedje. 2014. "Ribosomal Database Project: Data and Tools for High Throughput RRNA Analysis." *Nucleic Acids Research* 42 (Database issue). Oxford University Press: D633-42. https://doi.org/10.1093/nar/gkt1244.
Cox, Laura M, Shingo Yamanishi, Jiho Sohn, Alexander V Alekseyenko, Jacqueline M Leung, Ilseung Cho, Sungheon G Kim, et al. 2014. "Altering the Intestinal Microbiota during a Critical Developmental Window Has Lasting Metabolic Consequences." *Cell* 158 (4): 705–21. https://doi.org/10.1016/j.cell.2014.05.052.
David, Lawrence A., Corinne F. Maurice, Rachel N. Carmody, David B. Gootenberg, Julie E. Button, Benjamin E. Wolfe, Alisha V. Ling, et al. 2013. "Diet Rapidly and Reproducibly Alters the Human Gut Microbiome." *Nature* 505 (7484): 559–63. https://doi.org/10.1038/nature12820.
Davis, B M, E H Lawson, M Sandkvist, a Ali, S Sozhamannan, and M K Waldor. 2000. "Convergence of the Secretory Pathways for Cholera Toxin and the Filamentous Phage, CTXphi." *Science (New York, N.Y.)* 288 (April): 333–35. https://doi.org/10.1126/science.288.5464.333.
Deatherage, Daniel E, and Jeffrey E Barrick. 2014. "Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using Breseq." In , 165–88. https://doi.org/10.1007/978-1-4939-0554-6_12.
Dedrick, Rebekah M., Carlos A. Guerrero-Bustamante, Rebecca A. Garlena, Daniel A. Russell, Katrina Ford, Kathryn Harris, Kimberly C. Gilmour, et al. 2019. "Engineered Bacteriophages for Treatment of a Patient with a Disseminated Drug-Resistant Mycobacterium Abscessus." *Nature Medicine 2019 25:5* 25 (5). Nature Publishing Group: 730. https://doi.org/10.1038/s41591-019-0437-z.
Dethlefsen, Les, Sue Huse, Mitchell L Sogin, and David A Relman. 2008. "The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S RRNA Sequencing." *PLOS Biology* 6 (11): e280. https://doi.org/10.1371/journal.pbio.0060280.
Dethlefsen, Les, and David A Relman. 2011a. "Incomplete Recovery and Individualized Responses of the Human Distal Gut Microbiota to Repeated Antibiotic Perturbation." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl: 4554–61. https://doi.org/10.1073/pnas.1000087107.

Dodd, Dylan, Matthew H. Spitzer, William Van Treuren, Bryan D. Merrill, Andrew J. Hryckowian, Steven K. Higginbottom, Anthony Le, et al. 2017a. "A Gut Bacterial Pathway Metabolizes Aromatic Amino Acids into Nine Circulating Metabolites." *Nature* 551 (7682). Nature Publishing Group: 648. https://doi.org/10.1038/nature24661.

Doron, Shany, Sarah Melamed, Gal Ofir, Azita Leavitt, Anna Lopatina, Mai Keren, Gil Amitai, and Rotem Sorek. 2018. "Systematic Discovery of Antiphage Defense Systems in the Microbial Pangenome." *Science (New York, N.Y.)*, January. American Association for the Advancement of Science, eaar4120. https://doi.org/10.1126/science.aar4120.

Dorrestein, Pieter C, Sarkis K Mazmanian, and Rob Knight. 2014. "Finding the Missing Links among Metabolites, Microbes, and the Host." *Immunity* 40 (6): 824–32.

Duerkop, Breck A., Wenwen Huo, Pooja Bhardwaj, Kelli L. Palmer, and Lora V. Hooper. 2016. "Molecular Basis for Lytic Bacteriophage Resistance in Enterococci." *MBio* 7 (4). https://doi.org/10.1128/mBio.01304-16.

Duplessis, C., B. Biswas, B. Hanisch, M. Perkins, M. Henry, J. Quinones, D. Wolfe, L. Estrella, and T. Hamilton. 2017. "Refractory Pseudomonas Bacteremia in a 2-Year-Old Sterilized by Bacteriophage Therapy." *Journal of the Pediatric Infectious Diseases Society*, July. https://doi.org/10.1093/jpids/pix056.

Dy, Ron L., Corinna Richter, George P.C. Salmond, and Peter C. Fineran. 2014. "Remarkable Mechanisms in Microbes to Resist Phage Infections." *Annual Review of Virology* 1 (1). Annual Reviews : 307–31. https://doi.org/10.1146/annurev-virology-031413-085500.

Eren, A. Murat, Özcan C. Esen, Christopher Quince, Joseph H. Vineis, Hilary G. Morrison, Mitchell L. Sogin, and Tom O. Delmont. 2015. "Anvi'o: An Advanced Analysis and Visualization Platform for 'omics Data." *PeerJ* 3 (October). PeerJ Inc.: e1319. https://doi.org/10.7717/peerj.1319.

Estrella, Luis A, Javier Quinones, Matthew Henry, Ryan M Hannah, Robert K Pope, Theron Hamilton, Nimfa Teneza-Mora, Eric Hall, and Biswas Biswajit. 2016. "Characterization of Novel Staphylococcus Aureus Lytic Phage and Defining Their Combinatorial Virulence Using the OmniLog® System." *Bacteriophage* 6 (3). Taylor & Francis: e1219440. https://doi.org/10.1080/21597081.2016.1219440.

Faith, J. J., J. L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A. L. Goodman, J. C. Clemente, et al. 2013. "The Long-Term Stability of the Human Gut Microbiota." *Science* 341 (6141): 1237439–1237439. https://doi.org/10.1126/science.1237439.

Ferretti, Pamela, Edoardo Pasolli, Adrian Tett, Francesco Asnicar, Valentina Gorfer, Sabina Fedi, Federica Armanini, et al. 2018. "Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome." *Cell Host & Microbe* 24 (1). Cell Press: 133–145.e5. https://doi.org/10.1016/J.CHOM.2018.06.005.

Filippov, Andrey A., Kirill V. Sergueev, Yunxiu He, Xiao-Zhe Huang, Bryan T. Gnade, Allen J. Mueller, Carmen M. Fernandez-Prada, and Mikeljon P. Nikolich. 2011. "Bacteriophage-Resistant Mutants in Yersinia Pestis: Identification of Phage Receptors and Attenuation for Mice." Edited by Deepak Kaushal. *PLoS ONE* 6 (9). Public Library of Science: e25486. https://doi.org/10.1371/journal.pone.0025486.

Fitzgibbons, Shimae Cross, Yiming Ching, David Yu, Joe Carpenter, Michael Kenny, Christopher Weldon, Craig Lillehei, Clarissa Valim, Jeffrey D Horbar, and Tom Jaksic. 2009. "Mortality of Necrotizing Enterocolitis Expressed by Birth Weight Categories." *Journal of Pediatric Surgery* 44 (6): 1072–76.

https://doi.org/10.1016/j.jpedsurg.2009.02.013.
Flores, Cesar O, Justin R Meyer, Sergi Valverde, Lauren Farr, and Joshua S Weitz. 2011a. "Statistical Structure of Host-Phage Interactions." *Proceedings of the National Academy of Sciences of the United States of America* 108 (28). National Academy of Sciences: E288-97. https://doi.org/10.1073/pnas.1101595108.
Forde, Amanda, and Gerald F Fitzgerald. 1999. "Analysis of Exopolysaccharide (EPS) Production Mediated by the Bacteriophage Adsorption Blocking Plasmid, PCI658, Isolated from Lactococcus Lactis Ssp. Cremoris HO2." *International Dairy Journal* 9 (7). Elsevier: 465–72. https://doi.org/10.1016/S0958-6946(99)00115-6.
Frese, Steven A., Andra A. Hutton, Lindsey N. Contreras, Claire A. Shaw, Michelle C. Palumbo, Giorgio Casaburi, Gege Xu, et al. 2017. "Persistence of Supplemented *Bifidobacterium Longum* Subsp. *Infantis* EVC001 in Breastfed Infants." Edited by Rosa Krajmalnik-Brown. *MSphere* 2 (6). American Society for Microbiology Journals: e00501-17. https://doi.org/10.1128/mSphere.00501-17.
Fujimura, Kei E, Alexandra R Sitarik, Suzanne Havstad, Din L Lin, Sophia Levan, Douglas Fadrosh, Ariane R Panzer, et al. 2016. "Neonatal Gut Microbiota Associates with Childhood Multisensitized Atopy and T Cell Differentiation." *Nature Medicine* 22 (10): 1187–91. https://doi.org/10.1038/nm.4176.
Gelman, Daniel, Shaul Beyth, Vanda Lerer, Karen Adler, Ronit Poradosu-Cohen, Shunit Coppenhagen-Glazer, and Ronen Hazan. 2018. "Combined Bacteriophages and Antibiotics as an Efficient Therapy against VRE Enterococcus Faecalis in a Mouse Model." *Research in Microbiology*, May. Elsevier Masson. https://doi.org/10.1016/J.RESMIC.2018.04.008.
Gibson, Molly K, Bin Wang, Sara Ahmadi, Carey-Ann D Burnham, Phillip I Tarr, Barbara B Warner, and Gautam Dantas. 2016. "Developmental Dynamics of the Preterm Infant Gut Microbiota and Antibiotic Resistome." *Nature Microbiology* 1: 16024. https://doi.org/10.1038/nmicrobiol.2016.24.
Goodwin, Sara, Robert Wappel, and W. Richard McCombie. 2017. "1D Genome Sequencing on the Oxford Nanopore MinION." In *Current Protocols in Human Genetics*, 94:18.11.1-18.11.14. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/cphg.39.
Grier, Alex, Xing Qiu, Sanjukta Bandyopadhyay, Jeanne Holden-Wiltse, Haeja A. Kessler, Ann L. Gill, Brooke Hamilton, et al. 2017. "Impact of Prematurity and Nutrition on the Developing Gut Microbiome and Preterm Infant Growth." *Microbiome* 5 (1). BioMed Central: 158. https://doi.org/10.1186/s40168-017-0377-0.
Gu, Jingmin, Xiaohe Liu, Yue Li, Wenyu Han, Liancheng Lei, Yongjun Yang, Honglei Zhao, et al. 2012. "A Method for Generation Phage Cocktail with Great Therapeutic Potential." Edited by Brad Spellberg. *PLoS ONE* 7 (3). Public Library of Science: e31698. https://doi.org/10.1371/journal.pone.0031698.
Günthard, Huldrych F., Michael S. Saag, Constance A. Benson, Carlos del Rio, Joseph J. Eron, Joel E. Gallant, Jennifer F. Hoy, et al. 2016. "Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults." *JAMA* 316 (2). American Medical Association: 191. https://doi.org/10.1001/jama.2016.8900.
Hall, Alex R., Pauline D. Scanlan, Andrew D. Morgan, and Angus Buckling. 2011. "Host-Parasite Coevolutionary Arms Races Give Way to Fluctuating Selection." *Ecology Letters* 14 (7). Blackwell Publishing Ltd: 635–42. https://doi.org/10.1111/j.1461-0248.2011.01624.x.

Hanauer, David I, Mark J Graham, SEA-PHAGES, Laura Betancur, Aiyana Bobrownicki, Steven G Cresawn, Rebecca A Garlena, et al. 2017. "An Inclusive Research Education Community (IREC): Impact of the SEA-PHAGES Program on Research Outcomes and Student Learning." *Proceedings of the National Academy of Sciences of the United States of America* 114 (51). National Academy of Sciences: 13531–36. https://doi.org/10.1073/pnas.1718188115.

Hancock, Lynn E, and Michael S Gilmore. 2002. "The Capsular Polysaccharide of Enterococcus Faecalis and Its Relationship to Other Polysaccharides in the Cell Wall." *Proceedings of the National Academy of Sciences of the United States of America* 99 (3). National Academy of Sciences: 1574–79. https://doi.org/10.1073/pnas.032448299.

Harcombe, W R, and J J Bull. 2005. "Impact of Phages on Two-Species Bacterial Communities." *Applied and Environmental Microbiology* 71 (9). American Society for Microbiology (ASM): 5254–59. https://doi.org/10.1128/AEM.71.9.5254-5259.2005.

Heida, Fardou H, Van Zoonen, Anne G J F, Jan B F Hulscher, Te Kiefte, Bastiaan J C, Rianne Wessels, et al. n.d. "A Necrotizing Enterocolitis-Associated Gut Microbiota Is Present in the Meconium: Results of a Prospective Study" 62 (7): 863–70. Accessed March 10, 2017. https://doi.org/10.1093/cid/ciw016.

Hejnowicz, Monika S., Kamil Dąbrowski, Agnieszka Gozdek, Jarosław Kosakowski, Magdalena Witkowska, Magdalena I. Ulatowska, Beata Weber-Dąbrowska, et al. 2012. "Genomics of Staphylococcal Twort-like Phages - Potential Therapeutics of the Post-Antibiotic Era." *Advances in Virus Research* 83 (January). Academic Press: 143–216. https://doi.org/10.1016/B978-0-12-394438-2.00005-0.

Hendrickx, Antoni P A, Janetta Top, Jumamurat R Bayjanov, Hans Kemperman, Malbert R C Rogers, Fernanda L Paganelli, Marc J M Bonten, and Rob J L Willems. n.d. "Antibiotic-Driven Dysbiosis Mediates Intraluminal Agglutination and Alternative Segregation of Enterococcus Faecium from the Intestinal Epithelium." https://doi.org/10.1128/mBio.01346-15.

Henrick, Bethany M, Andra A Hutton, Michelle C Palumbo, Giorgio Casaburi, Ryan D Mitchell, Mark A Underwood, Jennifer T Smilowitz, and Steven A Frese. 2018. "Elevated Fecal PH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century." *MSphere* 3 (2). American Society for Microbiology Journals: e00041-18. https://doi.org/10.1128/mSphere.00041-18.

Herlemann, Daniel PR, Matthias Labrenz, Klaus Jürgens, Stefan Bertilsson, Joanna J Waniek, and Anders F Andersson. 2011. "Transitions in Bacterial Communities along the 2000 Km Salinity Gradient of the Baltic Sea." *The ISME Journal* 5 (10). Nature Publishing Group: 1571–79. https://doi.org/10.1038/ismej.2011.41.

Hesselbach, B A, and D Nakada. 1977. ""Host Shutoff" Function of Bacteriophage T7: Involvement of T7 Gene 2 and Gene 0.7 in the Inactivation of Escherichia Coli RNA Polymerase." *Journal of Virology* 24 (3): 736–45. http://www.ncbi.nlm.nih.gov/pubmed/338932.

Ho, Khang, Wenwen Huo, Savannah Pas, Ryan Dao, and Kelli L. Palmer. 2018. "Loss of Function Mutations in EpaR Confer Resistance to Phage NPV1 Infection in Enterococcus Faecalis OG1RF." *Antimicrobial Agents and Chemotherapy*, August. American Society for Microbiology, AAC.00758-18. https://doi.org/10.1128/AAC.00758-18.

Hodgson, D, L Shapiro, and K Amemiya. 1985. “Phosphorylation of the Beta’ Subunit of RNA Polymerase and Other Host Proteins upon PhiCd1 Infection of Caulobacter Crescentus.” *Journal of Virology* 55 (1). American Society for Microbiology (ASM): 238–41. http://www.ncbi.nlm.nih.gov/pubmed/16789252.

Hsiao, Elaine Y, Sara W McBride, Sophia Hsien, Gil Sharon, Embriette R Hyde, Tyler McCue, Julian A Codelli, et al. 2013. “Microbiota Modulate Behavioral and Physiological Abnormalities Associated with Neurodevelopmental Disorders.” *Cell* 155 (7). NIH Public Access: 1451–63. https://doi.org/10.1016/j.cell.2013.11.024.

Human Microbiome Project Consortium, The. 2012. “Structure, Function and Diversity of the Healthy Human Microbiome.” *Nature* 486. https://doi.org/10.1038/nature11234.

Jarvis, Audrey W., Lesley J. Collins, and H. W. Ackermann. 1993. “A Study of Five Bacteriophages of TheMyoviridae Family Which Replicate on Different Gram-Positive Bacteria.” *Archives of Virology* 133 (1–2). Springer-Verlag: 75–84. https://doi.org/10.1007/BF01309745.

Karav, Sercan, Annabelle Le Parc, Juliana Maria Leite Nobrega de Moura Bell, Steven A Frese, Nina Kirmiz, David E Block, Daniela Barile, and David A Mills. 2016. “Oligosaccharides Released from Milk Glycoproteins Are Selective Growth Substrates for Infant-Associated Bifidobacteria.” https://doi.org/10.1128/AEM.00547-16.

Keen, Eric C., Valery V. Bliskovsky, Francisco Malagon, James D. Baker, Jeffrey S. Prince, James S. Klaus, and Sankar L. Adhya. 2017. “Novel ‘Superspreader’ Bacteriophages Promote Horizontal Gene Transfer by Transformation.” Edited by Eduardo A. Groisman. *MBio* 8 (1). https://doi.org/10.1128/mBio.02115-16.

Kelly, Caleb J, Leon Zheng, Eric L Campbell, Bejan Saeedi, Carsten C Scholz, Amanda J Bayless, Kelly E Wilson, et al. 2015. “Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function.” *Cell Host & Microbe* 17 (5): 662–71. https://doi.org/10.1016/j.chom.2015.03.005.

Kelly, David, Olivia McAuliffe, Jim O’Mahony, and Aidan Coffey. 2011. “Development of a Broad-Host-Range Phage Cocktail for Biocontrol.” *Bioengineered Bugs* 2 (1). Taylor & Francis: 31–37. https://doi.org/10.4161/bbug.2.1.13657.

Keski-Nisula, Leea, Hanna-Reetta Kyynäräinen, Ulla Kärkkäinen, Jari Karhukorpi, Seppo Heinonen, and Juha Pekkanen. 2013. “Maternal Intrapartum Antibiotics and Decreased Vertical Transmission of Lactobacillus to Neonates during Birth.” *Acta Paediatrica (Oslo, Norway: 1992)* 102 (5): 480–85. https://doi.org/10.1111/apa.12186.

Khalifa, Leron, Yair Brosh, Daniel Gelman, Shunit Coppenhagen-Glazer, Shaul Beyth, Ronit Poradosu-Cohen, Yok-Ai Que, Nurit Beyth, and Ronen Hazan. 2015a. “Targeting Enterococcus Faecalis Biofilms with Phage Therapy.” *Applied and Environmental Microbiology* 81 (8). American Society for Microbiology: 2696–2705. https://doi.org/10.1128/AEM.00096-15.

Khalifa, Leron, Daniel Gelman, Mor Shlezinger, Axel Lionel Dessal, Shunit Coppenhagen-Glazer, Nurit Beyth, and Ronen Hazan. 2018. “Defeating Antibiotic- and Phage-Resistant Enterococcus Faecalis Using a Phage Cocktail in Vitro and in a Clot Model.” *Frontiers in Microbiology* 9 (February). Frontiers: 326. https://doi.org/10.3389/fmicb.2018.00326.

Khalifa, Leron, Mor Shlezinger, Shaul Beyth, Yael Houri-Haddad, Shunit Coppenhagen-Glazer, Nurit Beyth, and Ronen Hazan. 2016. “Phage Therapy against *Enterococcus Faecalis* in Dental Root Canals.” *Journal of Oral Microbiology* 8 (1). Taylor & Francis:

32157. https://doi.org/10.3402/jom.v8.32157.
Kim, Min-Soo, Eun-Jin Park, Seong Woon Roh, and Jin-Woo Bae. 2011. "Diversity and Abundance of Single-Stranded DNA Viruses in Human Feces." *Applied and Environmental Microbiology* 77 (22). American Society for Microbiology (ASM): 8062–70. https://doi.org/10.1128/AEM.06331-11.
Kind, Tobias, Hiroshi Tsugawa, Tomas Cajka, Yan Ma, Zijuan Lai, Sajjan S. Mehta, Gert Wohlgemuth, et al. 2017. "Identification of Small Molecules Using Accurate Mass MS/MS Search." *Mass Spectrometry Reviews*, April. https://doi.org/10.1002/mas.21535.
Kind, Tobias, Gert Wohlgemuth, Do Yup Lee, Yun Lu, Mine Palazoglu, Sevini Shahbaz, and Oliver Fiehn. 2009. "FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry." *Analytical Chemistry* 81 (24). American Chemical Society: 10038–48. https://doi.org/10.1021/ac9019522.
Klumpp, Jochen, Rob Lavigne, Martin J. Loessner, and Hans-Wolfgang Ackermann. 2010. "The SPO1-Related Bacteriophages." *Archives of Virology* 155 (10). Springer Vienna: 1547–61. https://doi.org/10.1007/s00705-010-0783-0.
Koch, Stefanie, Markus Hufnagel, Christian Theilacker, and Johannes Huebner. 2004. "Enterococcal Infections: Host Response, Therapeutic, and Prophylactic Possibilities." *Vaccine* 22 (7): 822–30. https://doi.org/10.1016/j.vaccine.2003.11.027.
Koenig, Jeremy E, Aymé Spor, Nicholas Scalfone, Ashwana D Fricker, Jesse Stombaugh, Rob Knight, Largus T Angenent, and Ruth E Ley. 2011. "Succession of Microbial Consortia in the Developing Infant Gut Microbiome." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 1 (Supplement 1). National Academy of Sciences: 4578–85. https://doi.org/10.1073/pnas.1000081107.
Kono, Nobuaki, and Kazuharu Arakawa. 2019. "Nanopore Sequencing: Review of Potential Applications in Functional Genomics." *Development, Growth & Differentiation*, April. John Wiley & Sons, Ltd (10.1111), dgd.12608. https://doi.org/10.1111/dgd.12608.
Kortright, Kaitlyn E., Benjamin K. Chan, Jonathan L. Koff, and Paul E. Turner. 2019a. "Phage Therapy: A Renewed Approach to Combat Antibiotic-Resistant Bacteria." *Cell Host & Microbe* 25 (2): 219–32. https://doi.org/10.1016/j.chom.2019.01.014.
Koskella, B., D. M. Lin, A. Buckling, and J. N. Thompson. 2012. "The Costs of Evolving Resistance in Heterogeneous Parasite Environments." *Proceedings of the Royal Society B: Biological Sciences* 279 (1735): 1896–1903. https://doi.org/10.1098/rspb.2011.2259.
Koskella, Britt, and Michael A Brockhurst. 2014. "Bacteria-Phage Coevolution as a Driver of Ecological and Evolutionary Processes in Microbial Communities." *FEMS Microbiology Reviews* 38 (5). Wiley-Blackwell: 916–31. https://doi.org/10.1111/1574-6976.12072.
Kostic, Aleksandar D, Dirk Gevers, Heli Siljander, Tommi Vatanen, Tuulia Hyötyläinen, Anu-Maaria Hämäläinen, Aleksandr Peet, et al. 2015. "The Dynamics of the Human Infant Gut Microbiome in Development and in Progression toward Type 1 Diabetes." *Cell Host & Microbe* 17 (2). Elsevier: 260–73. https://doi.org/10.1016/j.chom.2015.01.001.
Krug, Susanne, Gabi Kastenmüller, Ferdinand Stückler, Manuela J Rist, Thomas Skurk, Manuela Sailer, Johannes Raffler, et al. 2012. "The Dynamic Range of the Human Metabolome Revealed by Challenges." *The FASEB Journal* 26 (6): 2607–2619 %U http://www.fasebj.org/content/26/6/26.

Labrie, Simon J., Julie E. Samson, and Sylvain Moineau. 2010. “Bacteriophage Resistance Mechanisms.” *Nature Reviews Microbiology* 8 (5). Nature Publishing Group: 317–27. https://doi.org/10.1038/nrmicro2315.

Laguna, Theresa A, Cavan S Reilly, Cynthia B Williams, Cole Welchlin, and Chris H Wendt. 2015. “Metabolomics Analysis Identifies Novel Plasma Biomarkers of Cystic Fibrosis Pulmonary Exacerbation.” *Pediatric Pulmonology*, n/a-n/a %* © 2015 Wiley Periodicals, Inc. %U http:

Lamichhane, Santosh, Partho Sen, Alex M. Dickens, Matej Orešič, and Hanne Christine Bertram. 2018a. “Gut Metabolome Meets Microbiome: A Methodological Perspective to Understand the Relationship between Host and Microbe.” *Methods* 149 (October). Academic Press: 3–12. https://doi.org/10.1016/J.YMETH.2018.04.029.

Langmead, Ben, and Steven L Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59. https://doi.org/10.1038/nmeth.1923.

Larsen, Nadja, Finn K Vogensen, Frans W J van den Berg, Dennis Sandris Nielsen, Anne Sofie Andreasen, Bente K Pedersen, Waleed Abu Al-Soud, Søren J Sørensen, Lars H Hansen, and Mogens Jakobsen. 2010. “Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults.” *PLOS ONE* 5 (2): e9085. https://doi.org/10.1371/journal.pone.0009085.

Lawe-Davies, Olivia, and Simeon Bennett. 2017. “WHO Publishes List of Bacteria for Which New Antibiotics Are Urgently Needed.” Who. 2017. http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/.

Lebreton, François, Abigail L. Manson, Jose T. Saavedra, Timothy J. Straub, Ashlee M. Earl, and Michael S. Gilmore. 2017. “Tracing the Enterococci from Paleozoic Origins to the Hospital.” *Cell* 169 (5). Cell Press: 849–861.e13. https://doi.org/10.1016/J.CELL.2017.04.027.

Lebreton, Francois, Rob J. L. Willems, and Michael S. Gilmore. 2014. *Enterococcus Diversity, Origins in Nature, and Gut Colonization*. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary. http://www.ncbi.nlm.nih.gov/pubmed/24649513.

Lennon, Jay T, Sameed Ahmed M Khatana, Marcia F Marston, and Jennifer B H Martiny. 2007. “Is There a Cost of Virus Resistance in Marine Cyanobacteria?” *The ISME Journal* 1 (4). Nature Publishing Group: 300–312. https://doi.org/10.1038/ismej.2007.37.

Levi, Kyle, Mats Rynge, Eroma Abeysinghe, and Robert A. Edwards. 2018. “Searching the Sequence Read Archive Using Jetstream and Wrangler.” In *Proceedings of the Practice and Experience on Advanced Research Computing - PEARC ’18*, 1–7. New York, New York, USA: ACM Press. https://doi.org/10.1145/3219104.3229278.

Levy, Shawn E., and Richard M. Myers. 2016. “Advancements in Next-Generation Sequencing.” *Annual Review of Genomics and Human Genetics* 17 (1): 95–115. https://doi.org/10.1146/annurev-genom-083115-022413.

Liaw, Andy, and Matthew Wiener. 2002. “Classification and Regression by RandomForest.” *R News* 2 (3): 18–22.

Lim, Efrem S, Yanjiao Zhou, Guoyan Zhao, Irma K Bauer, Lindsay Droit, I Malick Ndao, Barbara B Warner, Phillip I Tarr, David Wang, and Lori R Holtz. 2015. “Early Life Dynamics of the Human Gut Virome and Bacterial Microbiome in Infants.” *Nature Medicine* 21 (10): 1228–34. https://doi.org/10.1038/nm.3950.

Lloyd-Price, Jason, Anup Mahurkar, Gholamali Rahnavard, Jonathan Crabtree, Joshua Orvis, A. Brantley Hall, Arthur Brady, et al. 2017. “Strains, Functions and Dynamics in the Expanded Human Microbiome Project.” *Nature* 550 (7674). Nature Research: 61. https://doi.org/10.1038/nature23889.

Lozupone, Catherine, and Rob Knight. 2005. “UniFrac: A New Phylogenetic Method for Comparing Microbial Communities.” *Applied and Environmental Microbiology* 71 (12). American Society for Microbiology: 8228–35. https://doi.org/10.1128/AEM.71.12.8228-8235.2005.

Lynch, Susan V, and Kenneth D Bruce. 2013. “The Cystic Fibrosis Airway Microbiome.” *Cold Spring Harbor Perspectives in Medicine* 3 (3 %U http://perspectivesinmedicine.cshlp.org/content/3/3/a009738).

Magill, Shelley S., Jonathan R. Edwards, Wendy Bamberg, Zintars G. Beldavs, Ghinwa Dumyati, Marion A. Kainer, Ruth Lynfield, et al. 2014. “Multistate Point-Prevalence Survey of Health Care–Associated Infections.” *New England Journal of Medicine* 370 (13). Massachusetts Medical Society : 1198–1208. https://doi.org/10.1056/NEJMoa1306801.

Mahé, Frédéric, Torbjørn Rognes, Christopher Quince, Colomban de Vargas, and Micah Dunthorn. 2014. “Swarm: Robust and Fast Clustering Method for Amplicon-Based Studies.” *PeerJ* 2 (September). PeerJ Inc.: e593. https://doi.org/10.7717/peerj.593.

Mailhammer, R, H L Yang, G Reiness, and G Zubay. 1975. “Effects of Bacteriophage T4-Induced Modification of Escherichia Coli RNA Polymerase on Gene Expression in Vitro.” *Proceedings of the National Academy of Sciences of the United States of America* 72 (12). National Academy of Sciences: 4928–32. http://www.ncbi.nlm.nih.gov/pubmed/1108008.

Manrique, Pilar, Benjamin Bolduc, Seth T Walk, John van der Oost, Willem M de Vos, and Mark J Young. 2016. “Healthy Human Gut Phageome.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (37). National Academy of Sciences: 10400–405. https://doi.org/10.1073/pnas.1601060113.

Marston, Marcia F, Francis J Pierciey, Alicia Shepard, Gary Gearin, Ji Qi, Chandri Yandava, Stephan C Schuster, Matthew R Henn, and Jennifer B H Martiny. 2012. “Rapid Diversification of Coevolving Marine Synechococcus and a Virus.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (12). National Academy of Sciences: 4544–49. https://doi.org/10.1073/pnas.1120310109.

Martiny, Jennifer B.H., Lasse Riemann, Marcia F. Marston, and Mathias Middelboe. 2014. “Antagonistic Coevolution of Marine Planktonic Viruses and Their Hosts.” *Annual Review of Marine Science* 6 (1). Annual Reviews : 393–414. https://doi.org/10.1146/annurev-marine-010213-135108.

Mcshan, William Michael, Eric Altermann, Paul Hyman, Alexa Ross, and Samantha Ward. 2016. “More Is Better: Selecting for Broad Host Range Bacteriophages.” https://doi.org/10.3389/fmicb.2016.01352.

Metsälä, J., A. Lundqvist, L. J. Virta, M. Kaila, M. Gissler, and S. M. Virtanen. 2015. “Prenatal and Post-Natal Exposure to Antibiotics and Risk of Asthma in Childhood.” *Clinical & Experimental Allergy* 45 (1): 137–45. https://doi.org/10.1111/cea.12356.

Meyer, Justin R, Devin T Dobias, Joshua S Weitz, Jeffrey E Barrick, R T Quick, and Richard E Lenski. 2012. “Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda.” *Science* 335 (6067): 428–32.

https://doi.org/10.1126/science.1214449.
Miller, T L. 1978. “The Pathway of Formation of Acetate and Succinate from Pyruvate by Bacteroides Succinogenes.” *Archives of Microbiology* 117 (2): 145–52. http://www.ncbi.nlm.nih.gov/pubmed/678020.
Minic, Zoran, Corinne Marie, Christine Delorme, Jean Michel Faurie, G??rald Mercier, Dusko Ehrlich, and Pierre Renault. 2007. “Control of EpsE, the Phosphoglycosyltransferase Initiating Exopolysaccharide Synthesis in Streptococcus Thermophilus, by EpsD Tyrosine Kinase.” *Journal of Bacteriology* 189 (4): 1351–57. https://doi.org/10.1128/JB.01122-06.
Minot, Samuel, Alexandra Bryson, Christel Chehoud, Gary D Wu, James D Lewis, and Frederic D Bushman. 2013. “Rapid Evolution of the Human Gut Virome.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (30). National Academy of Sciences: 12450–55. https://doi.org/10.1073/pnas.1300833110.
Minot, Samuel, Rohini Sinha, Jun Chen, Hongzhe Li, Sue A Keilbaugh, Gary D Wu, James D Lewis, and Frederic D Bushman. 2011. “The Human Gut Virome: Inter-Individual Variation and Dynamic Response to Diet.” *Genome Research* 21 (10). Cold Spring Harbor Laboratory Press: 1616–25. https://doi.org/10.1101/gr.122705.111.
Mizoguchi, Katsunori, Masatomo Morita, Curt R Fischer, Masatoshi Yoichi, Yasunori Tanji, and Hajime Unno. 2003a. “Coevolution of Bacteriophage PP01 and Escherichia Coli O157:H7 in Continuous Culture.” *Applied and Environmental Microbiology* 69 (1): 170–76. http://www.ncbi.nlm.nih.gov/pubmed/12513992.
Morona, R., L. Van den Bosch, and P. A. Manning. 1995. “Molecular, Genetic, and Topological Characterization of O-Antigen Chain Length Regulation in Shigella Flexneri.” *Journal of Bacteriology* 177 (4): 1059–68.
Morrison, Douglas J, and Tom Preston. 2016. “Formation of Short Chain Fatty Acids by the Gut Microbiota and Their Impact on Human Metabolism.” *Gut Microbes* 7 (3). Taylor & Francis: 189–200. https://doi.org/10.1080/19490976.2015.1134082.
Morrow, Ardythe L, Anne J Lagomarcino, Kurt R Schibler, Diana H Taft, Zhuoteng Yu, Bo Wang, Mekibib Altaye, et al. 2013. “Early Microbial and Metabolomic Signatures Predict Later Onset of Necrotizing Enterocolitis in Preterm Infants.” *Microbiome* 1 (1). BioMed Central: 13. https://doi.org/10.1186/2049-2618-1-13.
Mueller, N T, R Whyatt, L Hoepner, S Oberfield, M G Dominguez-Bello, E M Widen, A Hassoun, F Perera, and A Rundle. 2015. “Prenatal Exposure to Antibiotics, Cesarean Section and Risk of Childhood Obesity.” *International Journal of Obesity* 39 (4). Nature Publishing Group: 665–70. https://doi.org/10.1038/ijo.2014.180.
Nale, Janet, Tamsin Redgwell, Andrew Millard, and Martha Clokie. 2018a. “Efficacy of an Optimised Bacteriophage Cocktail to Clear Clostridium Difficile in a Batch Fermentation Model.” *Antibiotics* 7 (1). Multidisciplinary Digital Publishing Institute: 13. https://doi.org/10.3390/antibiotics7010013.
Nanthakumar, Nanda, Di Meng, Allan M Goldstein, Weishu Zhu, Lei Lu, Ricardo Uauy, Adolfo Llanos, Erika C Claud, and W Allan Walker. 2011. “The Mechanism of Excessive Intestinal Inflammation in Necrotizing Enterocolitis: An Immature Innate Immune Response.” *PLOS ONE* 6 (3): e17776. https://doi.org/10.1371/journal.pone.0017776.
Nechaev, Sergei, and Konstantin Severinov. 2003. “Bacteriophage-Induced Modifications of Host RNA Polymerase.” *Annual Review of Microbiology* 57 (1). Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA : 301–22.

https://doi.org/10.1146/annurev.micro.57.030502.090942.
Nikaido, Hiroshi. 2009. “Multidrug Resistance in Bacteria.” *Annual Review of Biochemistry* 78 (1). Annual Reviews : 119–46. https://doi.org/10.1146/annurev.biochem.78.082907.145923.
Nobel, Yael R, Laura M Cox, Francis F Kirigin, Nicholas A Bokulich, Shingo Yamanishi, Isabel Teitler, Jennifer Chung, et al. 2015. “Metabolic and Metagenomic Outcomes from Early-Life Pulsed Antibiotic Treatment.” *Nature Communications* 6: 7486. https://doi.org/10.1038/ncomms8486.
Oechslin, Frank, Philippe Piccardi, Stefano Mancini, Jérôme Gabard, Philippe Moreillon, José M Entenza, Gregory Resch, and Yok-Ai Que. 2016. “Synergistic Interaction between Phage Therapy and Antibiotics Clears Pseudomonas Aeruginosa Infection in Endocarditis and Reduces Virulence.” *Journal of Infectious Diseases*, December, jiw632-jiw632. https://doi.org/10.1093/infdis/jiw632.
Ogilvie, Lesley A, and Brian V Jones. 2015. “The Human Gut Virome: A Multifaceted Majority.” *Frontiers in Microbiology* 6. Frontiers Media SA: 918. https://doi.org/10.3389/fmicb.2015.00918.
Oksanen, Jari, F Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlinn, Peter R Minchin, et al. 2016. “Vegan: Community Ecology Package.” https://cran.r-project.org/package=vegan.
Oksanen, Jari, F Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R Minchin, R B O’Hara, Gavin L Simpson, Peter Solymos, M Henry H Stevens, and Helene Wagner. 2015. *Vegan: Community Ecology Package. R Package Version 2.3-5. 2016*.
Palmer, Chana, Elisabeth M Bik, Daniel B DiGiulio, David A Relman, and Patrick O Brown. 2007. “Development of the Human Infant Intestinal Microbiota.” Edited by Yijun Ruan. *PLoS Biology* 5 (7). Public Library of Science: e177. https://doi.org/10.1371/journal.pbio.0050177.
Palmer, K. L., P. Godfrey, A. Griggs, V. N. Kos, J. Zucker, C. Desjardins, G. Cerqueira, et al. 2012. “Comparative Genomics of Enterococci: Variation in Enterococcus Faecalis, Clade Structure in E. Faecium, and Defining Characteristics of E. Gallinarum and E. Casseliflavus.” *MBio* 3 (1): e00318-11. https://doi.org/10.1128/mBio.00318-11.
Pannaraj, Pia S, Fan Li, Chiara Cerini, Jeffrey M Bender, Shangxin Yang, Adrienne Rollie, Helty Adisetiyo, et al. 2017. “Association Between Breast Milk Bacterial Communities and Establishment and Development of the Infant Gut Microbiome.” *JAMA Pediatrics* 171 (7). American Medical Association: 647–54. https://doi.org/10.1001/jamapediatrics.2017.0378.
Paradis, E, J Claude, and K Strimmer. 2004. “A{PE}: Analyses of Phylogenetics and Evolution in {R} Language.” *Bioinformatics* 20: 289–90.
Paterson, Steve, Tom Vogwill, Angus Buckling, Rebecca Benmayor, Andrew J Spiers, Nicholas R Thomson, Mike Quail, et al. 2010. “Antagonistic Coevolution Accelerates Molecular Evolution.” *Nature* 464 (7286). Nature Publishing Group: 275–78. https://doi.org/10.1038/nature08798.
Pendleton, Jack N, Sean P Gorman, and Brendan F Gilmore. 2013. “Clinical Relevance of the ESKAPE Pathogens.” *Expert Review of Anti-Infective Therapy* 11 (3). Taylor & Francis: 297–308. https://doi.org/10.1586/eri.13.12.
Peng, Yu, Henry C M Leung, S M Yiu, and Francis Y L Chin. 2010. “IDBA – A Practical Iterative de Bruijn Graph De Novo Assembler.” In *Lecture Notes in Computer Science*

(Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6044 LNBI:426–40. https://doi.org/10.1007/978-3-642-12683-3_28.
Perry, Elizabeth B., Jeffrey E. Barrick, Brendan J. M. Bohannan, CC Traverse, MD Strand, and JJ Borges. 2015. "The Molecular and Genetic Basis of Repeatable Coevolution between Escherichia Coli and Bacteriophage T3 in a Laboratory Microcosm." Edited by Ramy K. Aziz. *PLOS ONE* 10 (6). Cambridge University Press: e0130639. https://doi.org/10.1371/journal.pone.0130639.
Pflughoeft, Kathryn J., and James Versalovic. 2012. "Human Microbiome in Health and Disease." *Annual Review of Pathology: Mechanisms of Disease* 7 (1). Annual Reviews : 99–122. https://doi.org/10.1146/annurev-pathol-011811-132421.
Pinto, Joana, M Rosário M Domingues, Eulália Galhano, Cristina Pita, Maria do Céu Almeida, Isabel M Carreira, and Ana M Gil. 2014. "Human Plasma Stability during Handling and Storage: Impact on NMR Metabolomics." *Analyst* 139 (5): 1168–77. https://doi.org/10.1039/C3AN02188B.
Price, Katherine E, Thomas H Hampton, Alex H Gifford, Emily L Dolben, Deborah A Hogan, Hilary G Morrison, Mitchell L Sogin, and George A O'Toole. 2013. "Unique Microbial Communities Persist in Individual Cystic Fibrosis Patients throughout a Clinical Exacerbation." *Microbiome* 1 (1): 27 %* 2013 Price et al.; licensee BioMed Central L.
Qin, Junjie, Ruiqiang Li, Jeroen Raes, Manimozhiyan Arumugam, Kristoffer Solvsten Burgdorf, Chaysavanh Manichanh, Trine Nielsen, et al. 2010. "A Human Gut Microbial Gene Catalogue Established by Metagenomic Sequencing." *Nature* 464 (7285). Nature Publishing Group: 59–65. https://doi.org/10.1038/nature08821.
Quinton, P M. 1983. "Chloride Impermeability in Cystic Fibrosis." *Nature* 301 (5899): 421–22. http://www.ncbi.nlm.nih.gov/pubmed/6823316.
Quinton, Paul M. 2008. "Cystic Fibrosis: Impaired Bicarbonate Secretion and Mucoviscidosis." *Lancet* 372 (9636): 415–17.
Raven, Kathy E., Sandra Reuter, Theodore Gouliouris, Rosy Reynolds, Julie E. Russell, Nicholas M. Brown, M. Estée Török, Julian Parkhill, and Sharon J. Peacock. 2016. "Genome-Based Characterization of Hospital-Adapted Enterococcus Faecalis Lineages." *Nature Microbiology* 1 (3). Nature Publishing Group: 15033. https://doi.org/10.1038/nmicrobiol.2015.33.
Reyes, Alejandro, Laura V Blanton, Song Cao, Guoyan Zhao, Mark Manary, Indi Trehan, Michelle I Smith, et al. 2015. "Gut DNA Viromes of Malawian Twins Discordant for Severe Acute Malnutrition." *Proceedings of the National Academy of Sciences of the United States of America* 112 (38). National Academy of Sciences: 11941–46. https://doi.org/10.1073/pnas.1514285112.
Rist, Manuela J, Claudia Muhle-Goll, Benjamin Görling, Achim Bub, Stefan Heissler, Bernhard Watzl, and Burkhard Luy. 2013. "Influence of Freezing and Storage Procedure on Human Urine Samples in NMR-Based Metabolomics." *Metabolites* 3 (2): 243–58. https://doi.org/10.3390/metabo3020243.
Robroeks, Charlotte M H H T, Joep J B N van Berkel, Jan W Dallinga, Quirijn Jöbsis, Luc J I Zimmermann, Han J E Hendriks, Miel F M Wouters, et al. 2010. "Metabolomics of Volatile Organic Compounds in Cystic Fibrosis Patients and Controls." *Pediatric Research* 68 (1): 75–80 %U http://www.ncbi.nlm.nih.gov/pubmed/203516.
Rosa, Patricio S La, Barbara B Warner, Yanjiao Zhou, George M Weinstock, Erica Sodergren, Carla M Hall-Moore, Harold J Stevens, et al. 2014. "Patterned Progression of Bacterial

Populations in the Premature Infant Gut." *Proceedings of the National Academy of Sciences* 111 (34): 12522–27. https://doi.org/10.1073/pnas.1409497111.
Roumpeka, Despoina D., R. John Wallace, Frank Escalettes, Ian Fotheringham, and Mick Watson. 2017. "A Review of Bioinformatics Tools for Bio-Prospecting from Metagenomic Sequence Data." *Frontiers in Genetics* 8 (March): 23. https://doi.org/10.3389/fgene.2017.00023.
Rutherford, Steven T, Courtney L Villers, Jeong-Hyun Lee, Wilma Ross, and Richard L Gourse. 2009. "Allosteric Control of Escherichia Coli RRNA Promoter Complexes by DksA." *Genes & Development* 23 (2): 236–48. https://doi.org/10.1101/gad.1745409.
Samson, Julie E., Alfonso H. Magadán, Mourad Sabri, and Sylvain Moineau. 2013a. "Revenge of the Phages: Defeating Bacterial Defences." *Nature Reviews Microbiology* 11 (10): 675–87. https://doi.org/10.1038/nrmicro3096.
Scanlan, Pauline D. 2017. "Bacteria–Bacteriophage Coevolution in the Human Gut: Implications for Microbial Diversity and Functionality." *Trends in Microbiology* 25 (8): 614–23. https://doi.org/10.1016/j.tim.2017.02.012.
Scanlan, Pauline D, Angus Buckling, and Alex R Hall. 2015. "Experimental Evolution and Bacterial Resistance: (Co)Evolutionary Costs and Trade-Offs as Opportunities in Phage Therapy Research." *Bacteriophage* 5 (2). Taylor & Francis: e1050153. https://doi.org/10.1080/21597081.2015.1050153.
Scanlan, Pauline D, Alex R Hall, Laura D C Lopez-Pascua, and Angus Buckling. 2011. "Genetic Basis of Infectivity Evolution in a Bacteriophage." *Molecular Ecology* 20 (5): 981–89. https://doi.org/10.1111/j.1365-294X.2010.04903.x.
Schmieder, R., and R. Edwards. 2011. "Quality Control and Preprocessing of Metagenomic Datasets." *Bioinformatics* 27 (6): 863–64. https://doi.org/10.1093/bioinformatics/btr026.
Schooley, Robert T, Biswajit Biswas, Jason J Gill, Adriana Hernandez-Morales, Jacob Lancaster, Lauren Lessor, Jeremy J Barr, et al. 2017. "Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails to Treat a Patient with a Disseminated Resistant Acinetobacter Baumannii Infection." *Antimicrobial Agents and Chemotherapy*, August. American Society for Microbiology, AAC.00954-17. https://doi.org/10.1128/AAC.00954-17.
Schrödinger, LLC. 2015. "The {PyMOL} Molecular Graphics System, Version~1.8."
Schulfer, Anjelique, and Martin J Blaser. 2015. "Risks of Antibiotic Exposures Early in Life on the Developing Microbiome." *PLOS Pathogens* 11 (7): e1004903. https://doi.org/10.1371/journal.ppat.1004903.
Sender, Ron, Shai Fuchs, and Ron Milo. 2016. "Revised Estimates for the Number of Human and Bacteria Cells in the Body." *PLoS Biology* 14 (8). Public Library of Science: e1002533. https://doi.org/10.1371/journal.pbio.1002533.
Sévin, Daniel C, Andreas Kuehne, Nicola Zamboni, and Uwe Sauer. 2015. "Biological Insights through Nontargeted Metabolomics." *Current Opinion in Biotechnology* 34 (August). Elsevier Current Trends: 1–8. https://doi.org/10.1016/J.COPBIO.2014.10.001.
Sharon, Itai, Michael J Morowitz, Brian C Thomas, Elizabeth K Costello, David A Relman, and Jillian F Banfield. 2013. "Time Series Community Genomics Analysis Reveals Rapid Shifts in Bacterial Species, Strains, and Phage during Infant Gut Colonization." *Genome Research* 23 (1). Cold Spring Harbor Laboratory Press: 111–20. https://doi.org/10.1101/gr.142315.112.

Silva, Juliano Bertozzi, Zachary Storms, and Dominic Sauvageau. 2016. “Host Receptors for Bacteriophage Adsorption.” *FEMS Microbiology Letters* 363. https://doi.org/10.1093/femsle/fnw002.

Sim, Kathleen, Alexander G Shaw, Paul Randell, Michael J Cox, Zoë E McClure, Ming-Shi Li, Munther Haddad, et al. 2015. “Dysbiosis Anticipating Necrotizing Enterocolitis in Very Premature Infants.” *Clinical Infectious Diseases* 60 (3): 389–97. https://doi.org/10.1093/cid/ciu822.

Singh, Kavindra V., Roshan J. Lewis, and Barbara E. Murray. 2009. “Importance of the *Epa* Locus of *Enterococcus Faecalis* OG1RF in a Mouse Model of Ascending Urinary Tract Infection.” *The Journal of Infectious Diseases* 200 (3): 417–20. https://doi.org/10.1086/600124.

Sjlund, Maria, Karin Wreiber, Dan I. Andersson, Martin J. Blaser, and Lars Engstrand. 2003. “Long-Term Persistence of Resistant *Enterococcus* Species after Antibiotics To Eradicate *Helicobacter Pylori*.” *Annals of Internal Medicine* 139 (6). American College of Physicians: 483. https://doi.org/10.7326/0003-4819-139-6-200309160-00011.

Smith, D. L., J. Dushoff, E. N. Perencevich, A. D. Harris, and S. A. Levin. 2004. “Persistent Colonization and the Spread of Antibiotic Resistance in Nosocomial Pathogens: Resistance Is a Regional Problem.” *Proceedings of the National Academy of Sciences* 101 (10): 3709–14. https://doi.org/10.1073/pnas.0400456101.

Smith, Daniel J, Alison C Badrick, Martha Zakrzewski, Lutz Krause, Scott C Bell, Gregory J Anderson, and David W Reid. 2014. “Pyrosequencing Reveals Transient Cystic Fibrosis Lung Microbiome Changes with Intravenous Antibiotics.” *European Respiratory Journal* 44 (4): 922–930 %U http://erj.ersjournals.com/content/44/4.

Sommer, Morten OA, and Gautam Dantas. 2011. “Antibiotics and the Resistant Microbiome.” *Current Opinion in Microbiology* 14 (5). Elsevier Current Trends: 556–63. https://doi.org/10.1016/J.MIB.2011.07.005.

Stern, Adi, Eran Mick, Itay Tirosh, Or Sagy, and Rotem Sorek. 2012. “CRISPR Targeting Reveals a Reservoir of Common Phages Associated with the Human Gut Microbiome.” *Genome Research* 22 (10). Cold Spring Harbor Laboratory Press: 1985–94. https://doi.org/10.1101/gr.138297.112.

Stewart, Christopher J, Nicholas D Embleton, Emma C L Marrs, Daniel P Smith, Andrew Nelson, Bashir Abdulkadir, Tom Skeath, et al. 2016. “Temporal Bacterial and Metabolic Development of the Preterm Gut Reveals Specific Signatures in Health and Disease.” *Microbiome* 4. https://doi.org/10.1186/s40168-016-0216-8.

Stingele, F, J R Neeser, B Mollet, Francesca Stingele, and Jean-richard Neeser. 1996. “Identification and Characterization of the Eps (Exopolysaccharide) Gene Cluster from Streptococcus Thermophilus Sfi6 . Identification and Characterization of the Eps (Exopolysaccharide) Gene Cluster from Streptococcus Thermophilus Sfi6.” *Journal of Bacteriology* 178 (6): 1680. https://doi.org/10.1128/jb.178.6.1680-1690.1996.

Stoddard, Lauren I, Jennifer B H Martiny, and Marcia F Marston. 2007. “Selection and Characterization of Cyanophage Resistance in Marine Synechococcus Strains.” *Applied and Environmental Microbiology* 73 (17). American Society for Microbiology: 5516–22. https://doi.org/10.1128/AEM.00356-07.

Stokholm, Jakob, Susanne Schjørring, Louise Pedersen, Anne Louise Bischoff, Nilofar Følsgaard, Charlotte G Carson, Bo L K Chawes, et al. 2013. “Prevalence and Predictors of Antibiotic Administration during Pregnancy and Birth.” *PloS One* 8 (12): e82932.

https://doi.org/10.1371/journal.pone.0082932.
Stoltz, David A, David K Meyerholz, and Michael J Welsh. 2015. “Origins of Cystic Fibrosis Lung Disease.” *New England Journal of Medicine* 372 (4): 351–362 %U http://dx.doi.org/10.1056/NEJMra1300109.
Stressmann, Franziska A, Geraint B Rogers, Peter Marsh, Andrew K Lilley, Thomas W V Daniels, Mary P Carroll, Lucas R Hoffman, et al. 2011. “Does Bacterial Density in Cystic Fibrosis Sputum Increase Prior to Pulmonary Exacerbation?” *Journal of Cystic Fibrosis: Official Journal of the European Cystic Fibrosis Society* 10 (5): 357–365 %U http://www.ncbi.nlm.nih.gov/pubmed/2166.
Summers, William C. 2012. “The Strange History of Phage Therapy.” *Bacteriophage* 2 (2). Taylor & Francis: 130. https://doi.org/10.4161/BACT.20757.
Tanaka, Shigemitsu, Takako Kobayashi, Prapa Songjinda, Atsushi Tateyama, Mina Tsubouchi, Chikako Kiyohara, Taro Shirakawa, Kenji Sonomoto, and Jiro Nakayama. 2009. “Influence of Antibiotic Exposure in the Early Postnatal Period on the Development of Intestinal Microbiota.” *FEMS Immunology and Medical Microbiology* 56 (1): 80–87. https://doi.org/10.1111/j.1574-695X.2009.00553.x.
Tate, S, G MacGregor, M Davis, J A Innes, and A P Greening. 2002. “Airways in Cystic Fibrosis Are Acidified: Detection by Exhaled Breath Condensate.” *Thorax* 57 (11): 926–929 %U http://www.ncbi.nlm.nih.gov/pubmed/1240.
Teng, F., K. V. Singh, A. Bourgogne, J. Zeng, and B. E. Murray. 2009. “Further Characterization of the Epa Gene Cluster and Epa Polysaccharides of Enterococcus Faecalis.” *Infection and Immunity* 77 (9): 3759–67. https://doi.org/10.1128/IAI.00149-09.
Teng, Fang, Kavindra V Singh, Agathe Bourgogne, Jing Zeng, and Barbara E Murray. 2009. “Further Characterization of the Epa Gene Cluster and Epa Polysaccharides of Enterococcus Faecalis.” *Infection and Immunity* 77 (9). American Society for Microbiology Journals: 3759–67. https://doi.org/10.1128/IAI.00149-09.
Tétart, F., F. Repoila, C. Monod, and H.M. Krisch. 1996. “Bacteriophage T4 Host Range Is Expanded by Duplications of a Small Domain of the Tail Fiber Adhesin.” *Journal of Molecular Biology* 258 (5): 726–31. https://doi.org/10.1006/jmbi.1996.0281.
Thaiss, Christoph A., Niv Zmora, Maayan Levy, and Eran Elinav. 2016. “The Microbiome and Innate Immunity.” *Nature* 535 (7610): 65–74. https://doi.org/10.1038/nature18847.
Thompson, John N. 1999. “Specific Hypotheses on the Geographic Mosaic of Coevolution.” *The American Naturalist* 153 (S5): S1–14. https://doi.org/10.1086/303208.
Thurlow, Lance R, Vinai Chittezham Thomas, and Lynn E Hancock. 2009. “Capsular Polysaccharide Production in Enterococcus Faecalis and Contribution of CpsF to Capsule Serospecificity.” *Journal of Bacteriology* 191 (20). American Society for Microbiology Journals: 6203–10. https://doi.org/10.1128/JB.00592-09.
Ting, Joseph Y, Anne Synnes, Ashley Roberts, Akhil Deshpandey, Kimberly Dow, Eugene W Yoon, Kyong-Soon Lee, Simon Dobson, Shoo K Lee, and Prakesh S Shah. 2016. “Association Between Antibiotic Use and Neonatal Mortality and Morbidities in Very Low-Birth-Weight Infants Without Culture-Proven Sepsis or Necrotizing Enterocolitis.” *JAMA Pediatrics* 170 (12): 1181–87. https://doi.org/10.1001/jamapediatrics.2016.2132.
Tornheim, Jeffrey A., and Kelly E. Dooley. 2019. “The Global Landscape of Tuberculosis Therapeutics.” *Annual Review of Medicine* 70 (1). Annual Reviews : 105–20.

https://doi.org/10.1146/annurev-med-040717-051150.
Turnbaugh, Peter J, Ruth E Ley, Michael A Mahowald, Vincent Magrini, Elaine R Mardis, and Jeffrey I Gordon. 2006. "An Obesity-Associated Gut Microbiome with Increased Capacity for Energy Harvest." *Nature* 444 (7122): 1027–1131. https://doi.org/10.1038/nature05414.
Twomey, Kate B, Mark Alston, Shi-Qi An, Oisin J O'Connell, Yvonne McCarthy, David Swarbreck, Melanie Febrer, J Maxwell Dow, Barry J Plant, and Robert P Ryan. 2013. "Microbiota and Metabolite Profiling Reveal Specific Alterations in Bacterial Community Structure and Environment in the Cystic Fibrosis Airway during Exacerbation." *PloS One* 8 (12): e82432.
Ubeda, Carles, Ying Taur, Robert R. Jenq, Michele J. Equinda, Tammy Son, Miriam Samstein, Agnes Viale, et al. 2010a. "Vancomycin-Resistant Enterococcus Domination of Intestinal Microbiota Is Enabled by Antibiotic Treatment in Mice and Precedes Bloodstream Invasion in Humans." *Journal of Clinical Investigation* 120 (12): 4332–41. https://doi.org/10.1172/JCI43918.
Uchiyama, Jumpei, Iyo Takemura, Miho Satoh, Shin-ichiro Kato, Takako Ujihara, Kazue Akechi, Shigenobu Matsuzaki, and Masanori Daibata. 2011. "Improved Adsorption of an Enterococcus Faecalis Bacteriophage ΦEF24C with a Spontaneous Point Mutation." Edited by Ramy K. Aziz. *PLoS ONE* 6 (10): e26648. https://doi.org/10.1371/journal.pone.0026648.
Underwood, Mark A., Jasmine C.C. Davis, Karen M. Kalanetra, Sanjay Gehlot, Sanjay Patole, Daniel J. Tancredi, David A. Mills, Carlito B. Lebrilla, and Karen Simmer. 2017. "Digestion of Human Milk Oligosaccharides by Bifidobacterium Breve in the Premature Infant." *Journal of Pediatric Gastroenterology and Nutrition*, April, 1. https://doi.org/10.1097/MPG.0000000000001590.
Van Tyne, Daria, and Michael S. Gilmore. 2014. "A Delicate Balance: Maintaining Mutualism to Prevent Disease." *Cell Host & Microbe* 16 (4). Cell Press: 425–27. https://doi.org/10.1016/J.CHOM.2014.09.019.
Venkataraman, Arvind, Miriam A. Rosenbaum, Sarah D. Perkins, Jeffrey J. Werner, and Largus T. Angenent. 2011. "Metabolite -Based Mutualism between Pseudomonas Aeruginosa PA14 and Enterobacter Aerogenes Enhances Current Generation in Bioelectrochemical Systems." *Energy & Environmental Science* 4 (11): 4550–59. https://doi.org/10.1039/C1EE01377G.
Venkataraman, Arvind, Miriam A Rosenbaum, Jeffrey J Werner, Stephen C Winans, and Largus T Angenent. 2014. "Metabolite Transfer with the Fermentation Product 2,3-Butanediol Enhances Virulence by Pseudomonas Aeruginosa %* © 2014 Nature Publishing Group %U Http://Www.Nature.Com/Ismej/Journal/Vaop/Ncurrent/Abs/Ismej2013232a.Html." *The ISME Journal*.
Viant, Mark R, Irwin J Kurland, Martin R Jones, and Warwick B Dunn. 2017. "How Close Are We to Complete Annotation of Metabolomes?" *Current Opinion in Chemical Biology* 36 (February). Elsevier Current Trends: 64–69. https://doi.org/10.1016/J.CBPA.2017.01.001.
Viertel, Tania Mareike, Klaus Ritter, and Hans Peter Horz. 2014. "Viruses versus Bacteria-Novel Approaches to Phage Therapy as a Tool against Multidrug-Resistant Pathogens." *Journal of Antimicrobial Chemotherapy* 69 (9): 2326–36.

https://doi.org/10.1093/jac/dku173.
Villarroel, Julia, Mette Voldby Larsen, Mogens Kilstrup, and Morten Nielsen. 2017. "Metagenomic Analysis of Therapeutic PYO Phage Cocktails from 1997 to 2014." *Viruses* 9 (11). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/v9110328.
Wandro, Stephen, Lisa Carmody, Tara Gallagher, John J LiPuma, and Katrine Whiteson. 2017. "Making It Last: Storage Time and Temperature Have Differential Impacts on Metabolite Profiles of Airway Samples from Cystic Fibrosis Patients." *MSystems* 2 (6). American Society for Microbiology Journals: e00100-17. https://doi.org/10.1128/mSystems.00100-17.
Wandro, Stephen, Andrew Oliver, Tara Gallagher, Claudia Weihe, Whitney England, Jennifer B. H. Martiny, and Katrine Whiteson. 2019. "Predictable Molecular Adaptation of Coevolving Enterococcus Faecium and Lytic Phage EfV12-Phi1." *Frontiers in Microbiology* 9 (January). Frontiers: 3192. https://doi.org/10.3389/fmicb.2018.03192.
Wandro, Stephen, Stephanie Osborne, Claudia Enriquez, Claudia Bixby, Antonio Arrieta, and Katrine Whiteson. 2017. "Microbial Community Assembly And Metabolite Profile Of The Gut Microbiome In Extremely Low Birthweight Infants." *BioRxiv*, April. Cold Spring Harbor Laboratory, 125922. https://doi.org/10.1101/125922.
Wang, Mingxun, Jeremy J Carver, Vanessa V Phelan, Laura M Sanchez, Neha Garg, Yao Peng, Don Duy Nguyen, et al. 2016. "Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking." *Nature Biotechnology* 34 (8). Nature Publishing Group: 828–37. https://doi.org/10.1038/nbt.3597.
Wattam, Alice R, James J Davis, Rida Assaf, Sébastien Boisvert, Thomas Brettin, Christopher Bun, Neal Conrad, et al. 2017. "Improvements to PATRIC, the All-Bacterial Bioinformatics Database and Analysis Resource Center." *Nucleic Acids Research* 45 (D1): D535–42. https://doi.org/10.1093/nar/gkw1017.
Whiteson, Katrine L, Simone Meinardi, Yan Wei Lim, Robert Schmieder, Heather Maughan, Robert Quinn, Donald R Blake, Douglas Conrad, and Forest Rohwer. 2014. "Breath Gas Metabolites and Bacterial Metagenomes from Cystic Fibrosis Airways Indicate Active PH Neutral 2,3-Butanedione Fermentation." *The ISME Journal* 8 (6): 1247–58.
Wickham, Hadley. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. http://ggplot2.org.
Wikoff, William R, Andrew T Anfora, Jun Liu, Peter G Schultz, Scott A Lesley, Eric C Peters, and Gary Siuzdak. 2009. "Metabolomics Analysis Reveals Large Effects of Gut Microflora on Mammalian Blood Metabolites." *Proceedings of the National Academy of Sciences of the United States of America* 106 (10). National Academy of Sciences: 3698–3703. https://doi.org/10.1073/pnas.0812874106.
Willemsen, L E M, M A Koetsier, S J H van Deventer, and E A F van Tol. 2003. "Short Chain Fatty Acids Stimulate Epithelial Mucin 2 Expression through Differential Effects on Prostaglandin E1 and E2 Production by Intestinal Myofibroblasts." *Gut* 52 (10): 1442–47. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1773837/.
Wright, A, C H Hawkins, E E Änggård, and D R Harper. 2009. "A Controlled Clinical Trial of a Therapeutic Bacteriophage Preparation in Chronic Otitis Due to Antibiotic-Resistant Pseudomonas Aeruginosa; A Preliminary Report of Efficacy." *Clinical Otolaryngology* 34 (4): 349–57. https://doi.org/10.1111/j.1749-4486.2009.01973.x.

Wright, Rosanna C. T., Ville-Petri Friman, Margaret C. M. Smith, and Michael A. Brockhurst. 2018. “Cross-Resistance Is Modular in Bacteria–phage Interactions.” Edited by Arjan de Visser. *PLOS Biology* 16 (10). Public Library of Science: e2006057. https://doi.org/10.1371/journal.pbio.2006057.

Yasbin, R E, V C Maino, and F E Young. 1976. “Bacteriophage Resistance in Bacillus Subtilis 168, W23, and Interstrain Transformants.” *Journal of Bacteriology* 125 (3). American Society for Microbiology (ASM): 1120–26. http://www.ncbi.nlm.nih.gov/pubmed/815237.

Yatsunenko, Tanya, Federico E Rey, Mark J Manary, Indi Trehan, Maria Gloria Dominguez-Bello, Monica Contreras, Magda Magris, et al. 2012. “Human Gut Microbiome Viewed across Age and Geography.” *Nature* 486 (7402): 222–27. https://doi.org/10.1038/nature11053.

Yen, Minmin, Lynne S. Cairns, and Andrew Camilli. 2017a. “A Cocktail of Three Virulent Bacteriophages Prevents Vibrio Cholerae Infection in Animal Models.” *Nature Communications* 8 (February): 14187. https://doi.org/10.1038/ncomms14187.

Yen, Sandi, Julie A K McDonald, Kathleen Schroeter, Kaitlyn Oliphant, Stanislav Sokolenko, Eric J M Blondeel, Emma Allen-Vercoe, and Marc G Aucoin. 2015. “Metabolomic Analysis of Human Fecal Microbiota: A Comparison of Feces-Derived Communities and Defined Mixed Communities.” *Journal of Proteome Research* 14 (3): 1472–82. https://doi.org/10.1021/pr5011247.

Yoong, Pauline, Raymond Schuch, Daniel Nelson, and Vincent A Fischetti. 2004a. “Identification of a Broadly Active Phage Lytic Enzyme with Lethal Activity against Antibiotic-Resistant Enterococcus Faecalis and Enterococcus Faecium.” *JOURNAL OF BACTERIOLOGY* 186 (14): 4808–12. https://doi.org/10.1128/JB.186.14.4808–4812.2004.

Zaman, Sojib Bin, Muhammed Awlad Hussain, Rachel Nye, Varshil Mehta, Kazi Taib Mamun, and Naznin Hossain. 2017. “A Review on Antibiotic Resistance: Alarm Bells Are Ringing.” *Cureus* 9 (6). Cureus Inc.: e1403. https://doi.org/10.7759/cureus.1403.

Zang, X, M E Monge, N A McCarty, A A Stecenko, and F M Fernández. 2017. “Feasibility of Early Detection of Cystic Fibrosis Acute Pulmonary Exacerbations by Exhaled Breath Condensate Metabolomics: A Pilot Study.” *J Proteome Res* 16 (2): 550–58. https://doi.org/10.1021/acs.jproteome.6b00675.

Zeng, M Y, N Inohara, and G Nuñez. 2017. “Mechanisms of Inflammation-Driven Bacterial Dysbiosis in the Gut.” *Mucosal Immunology* 10 (1). NIH Public Access: 18–26. https://doi.org/10.1038/mi.2016.75.

Zerbino, Daniel R, and Ewan Birney. 2008. “Velvet: Algorithms for de Novo Short Read Assembly Using de Bruijn Graphs.” *Genome Research* 18 (5): 821–29. https://doi.org/10.1101/gr.074492.107.

Zhao, J, J Li, P D Schloss, L M Kalikin, T A Raymond, J F Petrosino, V B Young, and J J LiPuma. 2011. “Effect of Sample Storage Conditions on Culture-Independent Bacterial Community Measures in Cystic Fibrosis Sputum Specimens.” *J Clin Microbiol* 49 (10): 3717–18. https://doi.org/10.1128/JCM.01189-11.

Zhao, Jiangchao, Charles R Evans, Lisa A Carmody, and John J LiPuma. 2015. “Impact of Storage Conditions on Metabolite Profiles of Sputum Samples from Persons with Cystic Fibrosis.” *Journal of Cystic Fibrosis: Official Journal of the European Cystic Fibrosis Society* 14 (4): 468–73.

Zhao, Jiangchao, Patrick D Schloss, Linda M Kalikin, Lisa A Carmody, Bridget K Foster, Joseph F Petrosino, James D Cavalcoli, et al. 2012. “Decade-Long Bacterial Community Dynamics in Cystic Fibrosis Airways.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (15): 5809–5814 %U http://www.ncbi.nlm.nih.gov/pubmed/22.

Zhvania, Pikria, Naomi Sulinger Hoyle, Lia Nadareishvili, Dea Nizharadze, and Mzia Kutateladze. 2017. “Phage Therapy in a 16-Year-Old Boy with Netherton Syndrome.” *Frontiers in Medicine* 4 (July). Frontiers: 94. https://doi.org/10.3389/fmed.2017.00094.