THE ROLE OF THE GUT MICROBIOME IN CHRONIC INFLAMMATORY DISEASES AND EVALUATION OF PHAGE FOR TARGETED MICROBIOME ENGINEERING

A Dissertation

by

JUSTIN XAVIER BOECKMAN

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Jason J. Gill
Committee Members,	Jan S. Suchodolski
	Thomas Matthew Taylor
	Ryland F. Young III
Head of Department,	Cliff G. Lamb

May 2022

Major Subject: Animal Science

ABSTRACT

The microbiome is closely linked with immune homeostasis and various disease states. Chronic inflammatory diseases such as environmental enteric dysfunction (EED) and inflammatory bowel disease (IBD) are associated with general gut microbiome dysbiosis. In addition to general intestinal microbiome dysbiosis, specific enteric microbes are also implicated in the initiation or maintenance of disease states, in this case pathogenic *Escherichia coli* and *Desulfovibrio* species. A targeted reduction of these microbes without additional contribution to overall dysbiosis may serve to alleviate symptoms of chronic inflammatory disease. Bacteriophages, or simply phages, are natural predators of bacteria and, given their narrow host ranges and lytic activity, are uniquely suited for targeted modulation of the intestinal microbiome.

This work describes the development of a piglet model for *E. coli*-induced gastroenteritis using a chronic repeated and onetime acute dosing regimen of pathogenic *E. coli* in an attempt to model EED. Acute and chronic dosing regimen of *E. coli*-induced similar responses in interleukin-6 and interlukin-8 as well as disruptions to the resident microbiome and metabolome. Next, bacteriophages were isolated against the trial strain in an attempt to alleviate *E. coli* induced clinical signs. While phage efficacy for the alleviation of piglet clinical signs was inconclusive, two novel phages, Minzhu and Mangalitsa, displaying preference for low-temperatures, were isolated that may be of use for biocontrol of *E. coli* for food safety applications.

Given *Desulfovibrio*'s putative role in IBD, we sought to isolate bacteriophage infecting *Desulfovibrio* and evaluate their potential for applied uses. In total, six phages infecting *D. desulfuricans* were isolated and characterized. Growth inhibition assays showed that *Desulfovibrio* phages such as these may be useful in future targeted reductions of *Desulfovibrio*. These phages were all highly novel and represent some of the only *Desulfovibrio* phages isolated, sequenced, and characterized to date.

Finally, as an extension of the exploration into *Desulfovibrio* phages, a functional prophage element within *D. desulfuricans* was identified. This prophage, ProddE, along with the related *Agrobacterium* phage Pasto described in this work, has broad sweeping implications for the evolution and lifestyle of the paradigm phage T7 and members of the *Autographiviridae*.

DEDICATION

Dedicated to my wife, Tracy, and son, Ryland, without whose love and support my PhD would not have been possible.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Jason Gill, for his guidance and support both inside and outside of the lab; and who always served as a wealth of knowledge and insights. I would like to thank my committee members Dr. Suchodolski, Dr. Taylor, and Dr. Young for their willingness to provide resources, insights, and guidance throughout my graduate experience.

I would also like to thank all members of the Gill lab past and present, Abby Korn, Season Xie, Denish Piya, Justin Leavitt, Jordyn Michalik, Daniel Mora, Miguel Gonzales, and Laura Acevedo for their friendship and support throughout my graduate career. I would particularly like to thank Abby Korn, Season Xie, and Denish Piya for taking me under their wing and showing me the ropes when I first arrived in the lab.

Finally, I would like to thank my entire immediate and extended family for their love and support. Particularly my wife, Tracy, who supported me through many late evenings and weekends in the lab; my son, Ryland, whose laughter and excitement kept me going during stressful periods; my loving parents, Mark and Kathy, who supported and encouraged me in all endeavors of life, particularly in scientific and academic pursuits; and my mother-in-law, Nancy, who has been a great support and resource for Tracy, Ryland and I.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was performed under the supervision of a dissertation committee consisting of Dr. Jason Gill (chair) and Dr. Matthew Taylor of the Animal Science Department, Dr. Jan Suchodolski of the College of Veterinary Medicine & Biomedical Sciences, and Dr. Ryland Young of the Biochemistry and Biophysics department.

The work performed in Chapter 2, the piglet animal model, was also performed in part by Sarah Sprayberry, whom is co-first author on the resulting publication. In particular, Sarah was the primary researcher in animal work involving housing, handling, tissue and fecal sample collection, and growth performance. Additionally, Dr. Abby Korn assisted in bacteriology and model development. Drs. Raquel Rech, Paula Giaretta, and Anna Blick performed animal necropsy. Drs. Kenneth Genovese and Todd Callaway provided animal housing and assistance in animal euthanasia. Dr. Jan Suchodolski provided qPCR resources, as well as microbiome and metabolomic insights. Drs. Chad Paulk and Jason Gill provided oversight and trial development regarding animal use and microbiology. I would like to thank Amanda Blake and So Young Park for assistance with qPCR. We would also like to thank Justin Leavitt, Jacob Chamblee, Jacob Lancaster, Lauren Lessor, Ruoyan Luo, Shayna Smith, Caitlin Older, Jeann Leal de Araujo, Kara Dunmire, Logan Joiner, and Lily Hernandez for their assistance with animal necropsy and sample preparation.

In chapter 3, the isolation and characterization of STEC phages, the annotation of Mangalitsa was performed by Cameron L Atkison as part of the Center for Phage Technology's Bacteriophage Genomics course.

In chapter 5, temperate T7-like phages and origins of *Autographiviridae*, phage Pasto and host strain was generously supplied by the lab of Dr. Carlos Gonzalez, through Abby Korn, Aravind Ravindran, and Guichan Yao. Tram Lee of the Center for Phage Technology prepared RNA from Pasto for sequencing. The Pasto genome was annotated aby Nikhil Manuel as part of the Center for Phage Technology's Bacteriophage Genomics course.

All other work described here was performed independently by the student.

Funding Sources

All work performed here was supported in part or in full by Texas A&M Agrilife Research, the Center for Phage Technology, and in the case of the Piglet model development, the Bill and Melinda Gate's foundation (OPP1139800).

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. Introduction and Literature Review

Initial observations of the human microbiome date all the way to the late 1600's, when the first observations and documentation of oral bacteria were made by Antonie van Leeuwenhoek with handmade microscopes (Leidy, 1853; Pariente, 2019). Later in the 1850's, a work entitled *A Flora and Fauna within Living Animals* was published by Joseph Leidy, considered to be foundational for the field of microbial research (Pariente, 2019). Approximately 40-50 years later, prominent microbiologists, such as Koch, Pasteur, Escherich, and Metchnikoff, famously recognized bacteria as the causative agents in a number of infectious diseases (Pariente, 2019). Beyond their recognition of pathogenic bacteria, these pioneers also appreciated a potential role for what would eventually be termed "commensal" bacteria in health, with Theodore Escherich, the eponym of the *Escherichia* genus, remarking: "…it would appear to be a pointless and doubtful exercise to examine and disentangle the apparently randomly appearing bacteria in normal feces and the intestinal tract, a situation that seems controlled by a thousand coincidences. If I have nevertheless devoted myself now for a year virtually exclusively to this special study, it was with the conviction that the accurate knowledge of these conditions is essential, for the understanding of not only the physiology of digestion, ... , but also the pathology and therapy of microbial intestinal diseases" (Pariente, 2019; Savage, 2001). The field of intestinal microbiota research received renewed vigor in the 1960's when multiple groups of investigators began work in this

field, while many of these findings were incomplete, they were able to establish a number of tenets of the intestinal microbiome such as: every individual has an intestinal microbial community composed of bacteria dominated by anaerobes such as *Bacteroides,* and levels of Bacteroides are several fold higher than species like *E. coli,* previously mischaracterized as dominating the microbiome (Savage, 2001).

More recently over the last two decades, advances in sequencing technology have led to vast leaps in our understanding of the intestinal microbiome. Nearly a decade ago, the Human Microbiome Project (HMP) reported one of the most comprehensive and large-scale studies of the healthy human microbiome from a variety of sampling sites including skin, vaginal, oral, and stool specimens. It was estimated the project identified 81-99% of the genera inhabiting the healthy western microbiome (Huttenhower et al., 2012). Studies such as the HMP and similar studies found the intestinal microbiome, consisting of all microbes within the gut, is dominated by bacterial and viral species (Huttenhower et al., 2012; Rinninella et al., 2019). The bacterial component of the microbiome has garnered the most investigation. The virome, consisting of both eukaryotic and prokaryotic viruses, has been comparatively understudied. This is likely due to lack of conserved features for the simplification of taxonomic determination as is possible for the bacterial genera. Additionally, the high ratio of gene content of unknown function has led to the virome's classification as "viral dark matter" (Aggarwala, Liang, & Bushman, 2017; Reyes, Semenkovich, Whiteson, Rohwer, & Gordon, 2012). Like the microbiome, changes of the understudied virome have also been associated with various intestinal disease states; whether these alterations

are contributors to disease or merely the consequence of shifting levels of available bacterial hosts within the gut remains to be determined. Given the various profiles of microbiome dysbiosis associated with gastrointestinal disease (Zheng, Liwinski, & Elinav, 2020), targeted approaches capable of modulating the intestinal microbiome back to a "healthy" state without significant harm to commensal bacteria (a feat not achievable with classic antibiotics) could represent novel treatment options for chronic intestinal conditions in which there are limit treatment options. This review aims to outline the current understanding of what constitutes a healthy microbiome and virome, the known dysbioses associated with select disease states, and explore how bacteriophage, or phage, may represent a unique tool for the targeted modulation the intestinal microbiome back to a state of healthy homeostasis.

1.2. What Constitutes a "Healthy" Microbiome?

The bacterial members of the human gut microbiome have received the most study to date, as bacteria are predominant members of the gut microbiota with an estimated 10^{13}-10^{14} bacteria in the colon (Sender, Fuchs, & Milo, 2016). Approximately 99% of genes within the gut have been reported to be of bacterial origin (Qin et al., 2010), and stool has been reported to have the greatest species richness of any sampled site on the body (Huse, Ye, Zhou, & Fodor, 2012). Given the term species "richness" has just been used, a brief definition of common metrics by which metagenomic sequencing data are frequently analyzed will be provided as they will be referenced heavily within this review. Firstly, "richness" is an indication of the number of different species within a sample, while "diversity" indicates how evenly and abundantly different taxa are

distributed (i.e., whether individual species dominate a microbiome or ecosystem). Additional details on these metrics can found in a recent review by Herath et al. (2017). As mentioned previously, the HMP reported one of the most comprehensive and large-scale studies of the healthy human microbiome approximately a decade ago, and it was estimated the project identified 81-99% of the genera inhabiting the healthy western microbiome (Huttenhower et al., 2012). The findings of the HMP, as well as the similar European Metagenomics of the Human Intestinal Tract project, MetaHIT (Qin et al., 2010), supported earlier small-scale studies of the intestine finding that the intestinal tract is dominated by two phyla, Bacteroidetes and Firmicutes; the phyla Actinobacteria, Proteobacteria, Fusobacteria, and Verrucomicrobia are also present but in lesser quantities (Arumugam et al., 2011; Huse et al., 2012; Huttenhower et al., 2012). Despite not being carried at relatively high abundance, the significance of these less abundant bacteria to the overall microbiome community cannot be ruled out as they can occupy specific niches or have been shown to contribute highly to abundant proteins in the microbiome (Arumugam et al., 2011). The outcomes of these studies, and similar studies, have demonstrated that there is a high degree of variation amongst an individual's bacteriome, and even operational taxonomy units (OTU's) that are represented in the majority of all subjects can have differences on the scale of orders of magnitudes in their relative carriage abundance (Huse et al., 2012; Huttenhower et al., 2012).

While there is a great deal of variation amongst taxa abundance between individuals, the abundance of metabolic pathways seems relatively stable (Huttenhower

et al., 2012). This suggests that the metabolic niches filled by the healthy microbiome within the gut may be more important than the exact microbial composition. Despite this variation in taxa abundance amongst individuals, there have been multiple attempts to classify various "enterotypes" or patterns of microbiota composition at the genus or species level detected by variations in abundance of an indicating taxon (Arumugam et al., 2011; Costea et al., 2018; Huse et al., 2012; J. Yang et al., 2020; Y. Yin et al., 2017). Although there is some ambiguity on whether these enterotypes may be more of a continuum versus discrete clusters or strata, there is general agreement on the existence of 3 distinct enterotypes: ET-B which is dominated by *Bacteroides,* ET-P dominated by *Prevotella,* and ET-R dominated by *Ruminococcus,* also at times referred to as ET-F for Firmicutes (Arumugam et al., 2011; Costea et al., 2018; Knights et al., 2014). Regardless of specific enterotype, the composition of an individual's microbiome has been shown to be relatively stable once an individual reaches adulthood (Lozupone, Stombaugh, Gordon, Jansson, & Knight, 2012). Faith et al. (2013) demonstrated 60% of an individual's bacteriome strains remained stable for 5 years; similar findings of microbiome stability over time have been repeatedly reported (Mehta et al., 2018; Y. Yin et al., 2017). Prior to adulthood, the microbiome of infants has been shown to be a reflection of the mother's microbiome dominated by *Bacteroides, Bifidobacterium, Parabacteroides,* and *Escherichia/Shigella,* but varies by type of delivery (vaginal vs caesarian) and breastfeeding status (breastfed vs bottle-fed) (Bäckhed et al., 2015). The microbiome of the newborn is also marked by low bacterial diversity, which rapidly increases with age before reaching its more stable "adult-like" state as early as 1-3 years

of age. However, it has also been reported bacterial diversity doesn't stabilize until an individual's twenties (Bäckhed et al., 2015; Efrem S. Lim et al., 2015; Odamaki et al., 2016) (Yatsunenko et al., 2012a). This transitionary period to an "adult-like" microbiome is associated with a reduction in Actinobacteria and Proteobacteria with a rise in Firmicutes (Odamaki et al., 2016). As the field has begun to understand the composition of the microbiome on macrolevel, more recent research has begun exploring the low abundance, uncultured, or unknown taxa of the intestinal tract by generating draft genomes from metagenomic sequencing (Stephen Nayfach, Shi, Seshadri, Pollard, & Kyrpides, 2019; Pasolli et al., 2019). Such exploration of the unknown members of the bacteriome will help to further our understanding of the bacteriome as a whole, since much of our ability to analyze the microbiome is based on comparison and read mapping to reference genomes.

The literature supports stable patterns to the intestinal microbiome, however there are a number of external factors that can influence our intestinal microbiomes. As mentioned above, age can play a large role in the composition of the microbiome with a succession of bacterial abundances from children, to adult, to the elderly (Efrem S. Lim et al., 2015; Odamaki et al., 2016; Stewart et al., 2018). Diet and culture can also play an important role in composition of microbiota as distinct geographic locations have distinct microbiomes (David et al., 2014; Lozupone et al., 2012; Yatsunenko et al., 2012b). Finally, both antibiotic and non-antibiotic drugs for disease treatment have been shown to affect the microbiome. While individual microbiomes can vary in response to different antibiotics, in general, antibiotics have been shown to have a rapid reduction in

microbiome diversity (Dethlefsen & Relman, 2011; Jakobsson et al., 2010). Once antibiotic treatment has ceased, microbiomes tend to trend back to their original state, although not always, as some have been reported to remain in a perturbed state for up to 4 years (until the end of the study) (Jakobsson et al., 2010). Additionally, these returns to the pre-antibiotic state are often only partial or incomplete (Dethlefsen & Relman, 2011). Similar prolonged antibiotic induced disruptions, as well as slow microbiome recovery, have also been reported in children (Korpela et al., 2016). Furthermore, it has been shown that once selected for, antibiotic resistance genes can persist in the microbiome for years at high levels after antibiotic treatment (Jakobsson et al., 2010). Outside of antibiotics, recent work has also examined the effect of non-antimicrobial therapeutic drugs on the intestinal microbiome. Studies have shown medications for the treatment of gastric acid disorders and antidiabetic drugs can lead to significant alterations of the intestinal microbial community (Forslund et al., 2015; Tsuda et al., 2015). In pursuit of this concept of non-antibiotic drug effects on the microbiome, Maier et al. screened over a 1,000 drugs on the market for antimicrobial activity against 40 intestine-associated bacterial strains and found that 24% of non-antibiotic drugs had some level of antimicrobial effects in-vitro (Maier et al., 2018), thus demonstrating yet another potential factor affecting our individual microbiomes.

As humans have co-evolved with their intestinal microbiota, the intestinal microbiome has also been repeatedly shown to play a critical role in modulation of the immune system and maintaining normal immune-homeostasis. This concept has recently been thoroughly reviewed by Zheng et al. (2020). But in brief, the presence of the

intestinal microbiome and specific commensals have been shown to play a key role in the formation of the intestine structural architecture and lymphoid tissue, in the expression of regulatory T cells, in B-cell development, and in the reduction of antibodies associated with allergic responses (Atarashi et al., 2013; Cahenzli, Köller, Wyss, Geuking, & McCoy, 2013; Gomez de Agüero et al., 2016; Mazmanian, Liu, Tzianabos, & Kasper, 2005; Wesemann et al., 2013). Prior and emerging research continues to suggest that our health and immune homeostasis are inextricably linked to the state of our microbiome. Thus, it is easy to see how disruptions to the microbiota could lead to chronic diseased states through aberrant activation of the immune system and unregulated inflammation.

Disease Associated Alterations to the Microbiome

Alterations to the healthy intestinal microbiome, or dysbiosis, have been associated with a number of chronic intestinal diseases; some have clear causative organisms, or combinations of organisms, while others remain more cryptic. A number of chronic diseases of the intestine such colon rectal cancer (CRC), recurrent *Clostridium difficile* infection, environmental enteric dysfunction (EED) and inflammatory bowel disease (IBD), have been associated with characteristic alterations to the microbiome. A summary of the alterations to the microbiome in these diseases can be found in **Figure 1, left.**

In brief, the microbiomes of CRC patients have been found to have significant increases in levels of *Bacteroides/Prevotella* genera, amongst other bacterial species, and decreased levels of butyrate-producing genera (Sobhani et al., 2011; Wang et al.,

2012). More recently, the tumor associated microenvironment has garnered interest, identifying specific bacterial species/strains with mechanisms for promoting tumor formation and/or growth such as pks+ *E. coli* and Enterotoxigenic *Bacteroides fragilis*, amongst others (Dejea et al., 2018; Drewes et al., 2017; Purcell, Visnovska, Biggs, Schmeier, & Frizelle, 2017; Saffarian et al., 2019; Yu et al., 2017). Furthermore, an intact commensal microbiota has been linked to positive cancer treatment outcomes, particularly with new immunotherapies (Gopalakrishnan et al., 2018; Matson et al., 2018; Routy et al., 2018) (N. Iida et al., 2013; Sivan et al., 2015; Vétizou et al., 2015; Viaud et al., 2013). Versus CRC wherein a number of specific microbes have been implicated in the CRC disease state, in persistent or recurrent *Clostridium difficile* infection (CDI) the causative organism of CDI is known and Koch's postulates have been displayed in an animal model (Bartlett, Onderdonk, Cisneros, & Kasper, 1977). However, the factors contributing to initial CDI and recurrence in ~20-30% patients (Dubberke & Olsen, 2012) are less clear, and novel treatments are areas of active research (Battaglioli et al., 2018). It has been shown that a perturbation of the resident microbiota by antibiotics or any diarrheal episode leading to a dysbiosis of the microbial community can leave individuals susceptible to CDIs and *C. difficile* blooms (Battaglioli et al., 2018; Brown Kevin, Khanafer, Daneman, & Fisman David, 2013; Buffie Charlie et al., 2012; Schäffler & Breitrück, 2018; VanInsberghe et al., 2020). Once CDI has taken hold, it has been shown that *C. difficile* toxin is capable of suppressing growth of Bacteroidaceae, an otherwise normally dominant member of the intestinal microbiome (Fletcher et al., 2021). Consistent with a level of dysbiosis responsible for maintaining

recurrent CDI, fecal microbiota transplants from healthy donors have been a promising treatment option to resolve recurrent CDI (Battaglioli et al., 2018).

Given EED and IBD are particularly relevant to Chapter 2, 3, and 4, the dysbiosis and implicated bacteria associated with these disease states will receive additional focus. Environmental Enteropathy, also known as environmental enteric dysfunction (EED), is a subclinical chronic inflammatory disorder primarily affecting developing regions where sanitization is not sufficient (K. Watanabe & W. A. Petri, 2016). EED is associated with blunting of small intestinal villi, intestinal inflammation, malnutrition, reduced efficacy of oral vaccine, and even reduced cognitive development, generally without overt diarrheal symptoms (Kosek et al., 2014; Naylor et al., 2015; K. Watanabe & W. A. Petri, 2016). It has been shown to be somewhat reversible and acquired by long term visitation or habitation in developing countries (Kosek et al., 2014; Lindenbaum, Gerson, & Kent, 1971). Nevertheless, some of its impacts such as stunting and growth faltering can have irreversible and long-lasting effects (Black et al., 2008). While the underlying cause of EED is currently unknown, it is believed to be due to chronic fecal-oral contamination leading to repeated exposure to low levels of fecal pathogens (Christine Marie George, Oldja, Biswas, Perin, Lee, Kosek, et al., 2015; Christine Marie George, Oldja, Biswas, Perin, Lee, Ahmed, et al., 2015; K. Watanabe & W. A. Petri, 2016). Early life exposure to enteropathogens, such as *Campylobacter, E. coli, Giardia* and *Cryptosporidium*, has been positively correlated with a number of EED biomarkers and symptoms like growth stunting (Iqbal et al., 2019a). These repeated exposures are believed to lead to chronic immune activation and increased gut permeability (K.

Watanabe & W. A. Petri, 2016). The malnutrition accompanying EED is also thought to leave the intestinal microbiome in an immature state with a reduction in bacterial diversity relative to healthy children, according to cohorts from Bangladesh and Malawi, whom also identified a number of potentially differential taxa, and notable reductions in the commensal bacteria *Bifidobacterium longum* and *Faecalibacterium prausnitzii* (Blanton Laura et al., 2016; Sathish Subramanian et al., 2014). It should be noted another metagenomic study utilizing 16S to explore the total bacteriome in EED found no significant changes in bacterial diversity and richness associated with growth rates of children with EED (Desai et al., 2020). Furthermore, it was found that therapeutic nutritional interventions were insufficient to bring the malnutrition gut microbiome into a healthy range (Sathish Subramanian et al., 2014). Additional interventions that have been attempted include zinc or multivitamins (Lauer et al., 2019), probiotics (Galpin et al., 2005), and antibiotics (Trehan, Shulman, Ou, Maleta, & Manary, 2009), amongst others; all have shown limited improvement, if any, illustrating the need for additional novel or combinatorial interventions. In support of a microbiome-dependent role for the malnourished aspect of EED, Blanton et al. (2016) also found transplantation of microbiota from healthy children into germ-free mice resulted in significant increase in weight gain compared to mice receiving microbiota transplants from malnourished children . Similarly, the hallmark shortening of intestinal villi has also been achieved in a mouse model by Brown et al. (2015) but only through a combination of a malnutrition diet and inoculation with a Bacteroidales and *E. coli* cocktail, suggesting a microbial and diet synergism contributing to EED. A recent study by Desai et al. (2020) found there

existed significant differences in abundance of 30 taxa in EED children that exhibited poor growth. Furthermore, 10 of these taxa frequently associated with healthy gut homeostasis were completely undetected in children exhibiting poor growth, including species from *Megasphaera*, *Prevotella*, *Bacteroides*, and *Bifidobacterium* genera (Desai et al., 2020). Thus, while the exact cause of EED remains enigmatic, its onset is most likely a consequence of the combination of diet, geographic location, fecal pathogen exposure, and alterations of the intestinal microbiome. All intervention attempts to date have only shown limited improvement, if any, suggesting alternative novel interventions or novel combinatorial treatments are necessary to alleviate EED symptoms.

Finally, IBD is a chronic inflammatory disease characterized by inflammation of the gastrointestinal tract. IBD is comprised of two similar yet distinct disorders, Crohn's disease (CD) and ulcerative colitis (UC) (M. Zhang et al., 2017). While IBD has been extensively studied, the underlying cause remains unknown. Prevailing theory suggests an intestinal dysbiosis of the microbiota leads to a disruption of immune homeostasis and thus a chronic aberrant inflammatory response (M. Zhang et al., 2017). Like EED, no singular causative organism has yet been identified. In recent years, studies have solidified this connection between intestinal microbiome dysbiosis and IBD. Patients with IBD have been shown to have an altered intestinal microbiome with lower bacterial diversity and abundance relative to controls, and inoculation of this IBD associated microbiota into germ free mice induced abnormalities in adaptive immunity (Torres et al., 2020). Similarly, transplantation of IBD associated microbiota into mice has been repeatedly shown to induce IBD-like ileitis and/or proinflammatory responses, including

increases in proinflammatory T cells such as Th17 and Th2 (Britton et al., 2019; Schaubeck et al., 2016). Thus, it is clear that the microbiota plays a pivotal role in the pathogenesis of IBD; thus far, however, Koch's postulates have yet to be demonstrated with a single organism. Moreover, comparable disease states have only been achieved by inoculation of a microbial consortia. Despite this, a number of -omic studies have attempted to elucidate specific bacteria involved in the development of IBD. Metagenomics from IBD patients show a general reduction in microbial diversity, expansion of *E. coli* populations, enrichment of *Ruminococcus gnavus*, and a reduction in *Faecalibacterium prasniitzii* and *Roseburia hominis* (Franzosa et al., 2019; Hall et al., 2017; Lloyd-Price et al., 2019; Ryan et al., 2020). However, these studies often result in a large number of differential species up to 50 in some cases (Franzosa et al., 2019). This makes attribution of causal mechanism(s) to individual species impractical. Furthermore, when coupled with transcriptomics, certain *Clostridium* spp. display an increase in relative transcription levels disproportionate to genome abundance during periods of IBD dysbiosis, implying under-represented species may also contribute to IBD (Lloyd-Price et al., 2019). Additionally, organisms such as *Helicobacter pylori,* known to be associated with peptic ulcers, have been suggested as an initiator of IBD (Mansour et al., 2018); however, other work suggests *H. pylori* has a negative association of with IBD and perhaps even provides a protective effect (Ding et al., 2021; Homolak et al., 2021). These conflicting results leave the role of *H. pylori* in IBD, if any, questionable. Of note, the sulfate-reducing bacterial genus *Desulfovibrio* has also been thought to contribute to initiation and maintenance of IBD (Gibson, Cummings, &

Macfarlane, 1991a) (Zinkevich & Beech, 2000b). While present in both healthy and diseased individuals with IBD, multiple studies have reported higher incidences/colonization of SRB as well as increased levels of hydrogen sulfide gas (H_2S), a byproduct of the *Desulfovibrio* sulfate reduction pathway, in those diagnosed with IBD (Gibson et al., 1991a) (Coutinho et al., 2017a; Fite, 2004b; Levine, Ellis, Furne, Springfield, & Levitt, 1998a; Loubinoux, Bronowicki, Pereira, Mougenel, & Le Faou, 2002b; Pitcher, Beatty, & Cummings, 2000a; Zinkevich & Beech, 2000b). The H_2S produced by *Desulfovibrio* is a plausible mechanism for contributing to the state of dysbiosis associated with IBD, given the gas is cytotoxic and inhibits the oxidation of the important short chain fatty acid butyrate (Singh & Lin, 2015b). However as with most of the bacterial species implicated with causative roles of IBD, *Desulfovibrio* requires more study to cement links between it and IBD.

In summary, there is no single "healthy" microbiome, but rather a continuum of unique healthy microbiomes which follow general patterns of taxon abundance and carriage. Moreover, the composition of the microbiome is shaped by various factors, like age, diet, culture, and antibiotic or therapeutic drug use. Deviations from these healthy microbiome trends, or dysbioses, are often marked by reductions in bacterial diversity, and can result in a variety of diseased states, though in many cases it is still unclear if disease-associated alterations of the microbiome are the cause or the consequence of the disease state. A summary of healthy and dysbiotic microbiomes can be seen in **Figure 1, left.** Given these disease-associated dysbioses, a tool for targeted modulation of the

microbiome back to a “healthy” state without further disruption of the resident microbiome may alleviate the symptoms associated with dysbiotic disease states.

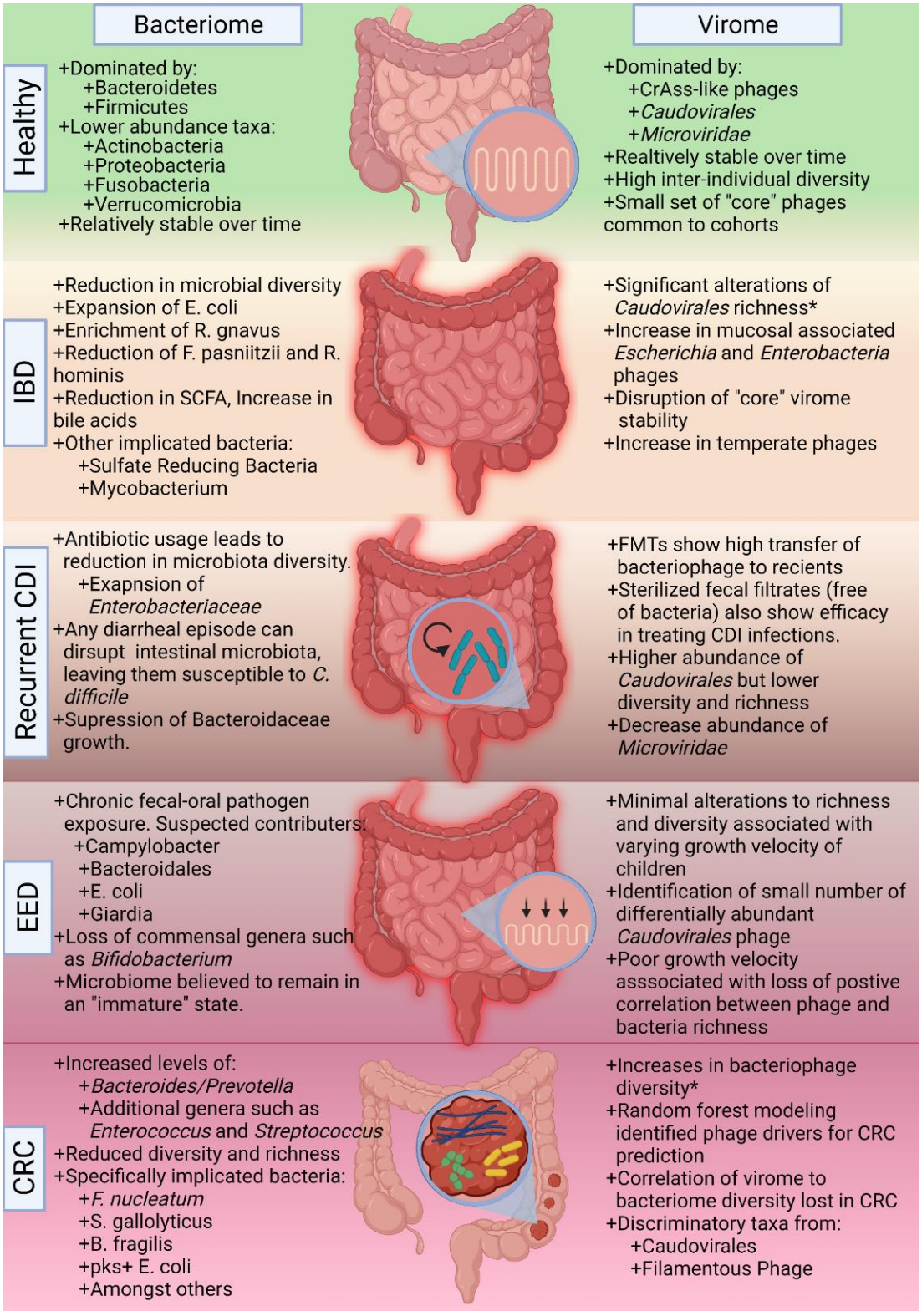

Bacteriome
Virome
Healthy
+Dominated by:
+Bacteroidetes
+Firmicutes
+Lower abundance taxa:
+Actinobacteria
+Proteobacteria
+Fusobacteria
+Verrucomicrobia
+Relatively stable over time
+Dominated by:
+CrAss-like phages
+Caudovirales
+Microviridae
+Realtively stable over time
+High inter-individual diversity
+Small set of "core" phages common to cohorts
IBD
+Reduction in microbial diversity
+Expansion of E. coli
+Enrichment of R. gnavus
+Reduction of F. pasniitzii and R. hominis
+Reduction in SCFA, Increase in bile acids
+Other implicated bacteria:
+Sulfate Reducing Bacteria
+Mycobacterium
+Significant alterations of Caudovirales richness*
+Increase in mucosal associated Escherichia and Enterobacteria phages
+Disruption of "core" virome stability
+Increase in temperate phages
Recurrent CDI
+Antibiotic usage leads to reduction in microbiota diversity.
+Exapnsion of Enterobacteriaceae
+Any diarrheal episode can dirsupt intestinal microbiota, leaving them susceptible to C. difficile
+Supression of Bacteroidaceae growth.
+FMTs show high transfer of bacteriophage to recients
+Sterilized fecal filtrates (free of bacteria) also show efficacy in treating CDI infections.
+Higher abundance of Caudovirales but lower diversity and richness
+Decrease abundance of Microviridae
EED
+Chronic fecal-oral pathogen exposure. Suspected contributers:
+Campylobacter
+Bacteroidales
+E. coli
+Giardia
+Loss of commensal genera such as Bifidobacterium
+Microbiome believed to remain in an "immature" state.
+Minimal alterations to richness and diversity associated with varying growth velocity of children
+Identification of small number of differentially abundant Caudovirales phage
+Poor growth velocity asssociated with loss of postive correlation between phage and bacteria richness
CRC
+Increased levels of:
+Bacteroides/Prevotella
+Additional genera such as Enterococcus and Streptococcus
+Reduced diversity and richness
+Specifically implicated bacteria:
+F. nucleatum
+S. gallolyticus
+B. fragilis
+pks+ E. coli
+Amongst others
+Increases in bacteriophage diversity*
+Random forest modeling identified phage drivers for CRC prediction
+Correlation of virome to bacteriome diversity lost in CRC
+Discriminatory taxa from:
+Caudovirales
+Filamentous Phage

Figure 1 Summary of General Trends of Microbiome and Virome in Various Diseased States

General trends of the microbiome (left) and virome (right) in healthy and various diseased states. (Center) Intestine diagram with a hallmark symptom associated disease. Healthy: No inflammation and healthy finger like villi. Inflammatory Bowel Disease (IBD): general inflammation of intestine, localization of inflammation varies between ulcerative colitis and Crohn's disease. Recurrent *Clostridium difficile* infection (CDI): inflammation and chronic expansion of *C. difficile*. Environmental Enteric Dysfunction (EED): blunting of intestinal villi and malnutrition. Colorectal Cancer (CRC): Tumor development, with tumor microenvironment colonized by pro-oncogenic bacterial species. See associated diseased state text for additional details and supporting literature. *Indicates a trend in which literature has not yet reached consensus or a opposite result has also been reported within literature.

1.3. Bacteriophage and the "Healthy Virome"

Bacteriophages

Bacteriophages, or simply phages, may represent the perfect tool for this difficult task of modulating the microbiome with minimal collateral damage to off-target bacterial species. Independently discovered in 1915 and 1917 by Frederick William Twort and Felix d'Herelle, respectively, phages are viruses of bacteria. Phages were recognized very early on for their potential to efficiently kill bacteria (Clokie, Millard, Letarov, & Heaphy, 2011; Gill & Hyman, 2010a). After their initial discovery, phages experienced a brief heyday of therapeutic and prophylactic applications as anti-bacterial

agents(Fruciano & Bourne, 2007; Gill & Hyman, 2010a). However, due to limited understandings of phage biology at the time and the rise of antibiotic use, therapeutic applications of phage waned in much of the Western world, although the use of phage persisted in eastern European and former Soviet countries (Fruciano & Bourne, 2007; Gill & Hyman, 2010a). Recently, with the rise of antibiotic resistant bacteria, phage have received renewed interest as antimicrobials, given their narrow host range relative to antibiotics, and potent lytic activity. In line with this renewed interest, phages have been granted the status of Generally Recognized as Safe (GRAS) as food additives (GRAS Notice No. GRN 000603) and commercial formulations have been used to control foodborne pathogens (Kazi & Annapure, 2016). The narrow host range and potent lytic activity of phages have made them attractive options for therapeutic treatment of acute infections (Gill & Hyman, 2010a; Schooley et al., 2017b; Young & Gill, 2015a). Specifically, the recent high-profile use of phage to treat a life-threatening antibiotic resistant *Acinetobacter baumannii* infection garnered attention (Schooley et al., 2017b). Following positive patient outcomes from compassionate or emergency use cases, the FDA has recently approved clinical trials for the intravenous administration of phages (Voelker, 2019), and a number of companies focused on the development of phage therapy use in humans have formed. In addition to the use of phage for treating acute infections, here we address whether bacteriophage may be novel and effective tools for modulating the intestinal microbiome and resolving disease-associated dysbiosis.

Bacteriophages are estimated to be the most numerous organisms on the planet with greater than 10^{30} tailed phages in the biosphere (Brüssow & Hendrix, 2002).

Caudovirales, or the tailed phages, are the most heavily studied group of phages and are a diverse order of dsDNA viruses consisting of 14 families and 73 subfamilies (Lefkowitz et al., 2018). Beyond the order *Caudovirales*, other relevant taxa in the context of the intestinal virome are bacteriophages of the family *Microviridae* and the family *Inoviridae*. The family *Microviridae* are extremely small both in terms of virion and genome size with ~5kb ssDNA viruses with icosahedral capsids. In contrast, family *Inoviridae* are filamentous ssDNA bacteriophages with notably different lifestyles compared to *Caudovirales* and *Microviridae* in that *Inoviridae* do not lyse or kill the host cell to release progeny; virions are instead continuously secreted from infected cells (Knezevic, Adriaenssens, & Ictv Report, 2021; Lefkowitz et al., 2018). *Caudovirales* phage can be categorized into two lifestyles, either the virulent lifestyle or the temperate lifestyle (Gill & Hyman, 2010a). Virulent phages are restricted to the lytic pathway and follow the generalized lifecycle shown in **Figure 2A** to lyse their bacterial hosts. The lytic cycle can be broken into three major steps: 1) Infection: Phage virions adsorb to a host bacterial cell by recognition of its cognate phage receptor, and phage DNA is transferred into the bacterial cell. This binding by the phage tail fiber, often to an outer membrane protein or LPS component on the surface of the bacterial host, is the initial determinant of the highly specific and narrow phage host range. 2) Replication: the phage takes over the host cell using various cellular components for its DNA replication and production of phage structural protein, resulting in the assembly of new virions 3) Lysis: phage proteins disrupt the cell envelope, resulting in the explosive lysis of the bacterial cell releasing new phage progeny. Detailed reviews of phage adsorption, virion

assembly, and lysis are available (Aksyuk & Rossmann, 2011; Cahill & Young, 2019; Dowah & Clokie, 2018; Rakhuba, Kolomiets, Dey, & Novik, 2010). While virulent phages are obligately lytic and only capable of host cell lysis **(Figure 2A)**, phages classified as temperate are capable of entering both the lytic cycle as well as the lysogenic cycle. At the time of initial infection, temperate phages make a decision to either enter the lytic or lysogenic cycle. In the lysogenic cycle, phages are integrated into the host chromosome or maintained as an extrachromosomal element and peacefully maintained for a time **(Figure 2B)** (Gill & Hyman, 2010a). These peacefully maintained phages are termed prophages, and the bacterial hosts that they occupy are termed lysogens (Gill & Hyman, 2010a). During the lysogenic cycle (**Figure 2B**) temperate phages perform the following generalized steps: 4) Phage DNA is integrated into the host chromosome turning the host into a lysogen. 5) Prophages while integrated into the host chromosome remain repressed and replicate passively along with the division of the host cell. 6) A specific induction signal is met, commonly the host SOS response, and integrated phage elements excise and proceed to the lytic cycle (Step 2, **Figure 2A**) (Gill & Hyman, 2010a).

Bacteriophages are predicted to be drivers of bacterial evolution (and vice versa) through the exertion of selection pressure leading to the oft mentioned “arms race” between phage and the evolving anti-phage defense systems. Prime examples of this evolutionary arms race can be seen from the famous bacterial CRISPR-Cas system and identification of phage-encoded anti-CRISPR proteins (Landsberger et al., 2018), or phage encoded anti-restriction systems, such as the (defense against restriction) DarAB

system (S. Iida, Streiff, Bickle, & Arber, 1987; Piya, Vara, Russell, Young, & Gill, 2017). These examples are only a fraction of the elegantly intertwined defense and counter-defense systems produced through eons of co-evolution. Additional bacterial defenses against phage include surface receptor modification, abortive infection, or inhibition of phage assembly; such systems have been reviewed by Seed (2015). Depending on the ecological or biological context in which phage-host interactions take place, phages are frequently described as following "Kill-the-winner", "red-queen", or "piggy-back the winner" models of predator/prey dynamic. These concepts describe phages' control of bacterial populations and co-evolution without leading to the total extinction of either the predator (phage) or the prey (bacteria). Thorough reviews of these phage-host dynamic models are available (Avrani, Schwartz, & Lindell, 2012; Andrey N Shkoporov & Hill, 2019; Silveira & Rohwer, 2016). In addition to the purely antagonistic shaping of bacterial communities through the phage lytic cycle, the lysogenic cycle of temperate phages offers an additional layer of altering bacterial communities. Once integrated into the bacterial host, an individual phage's fitness is closely tied to the fate of the bacterial host. As such, temperate phages often have "morons", segments of DNA expressing accessory proteins that alter the phenotype and fitness of the bacterial host cell (Cumby, Davidson, & Maxwell, 2012; Taylor, Fitzpatrick, Islam, & Maxwell, 2019; Tsao et al.). Phage morons have the ability to make the lysogenized host more resistant to infection by other phages, display increased virulence through carriage of toxins (famously Shiga-toxin), modulate host motility, and increase resistance to antibiotics (Taylor et al., 2019; Tsao et al.). In summary,

bacteriophages are capable of exerting control over microbial populations through phage-mediated killing, driving bacterial evolution through the evolutionary arms race, and altering the fitness of bacterial hosts through the formation of lysogens. These natural modulatory effects of phage make them ideal candidates for targeted modulation of microbial populations in the intestine.

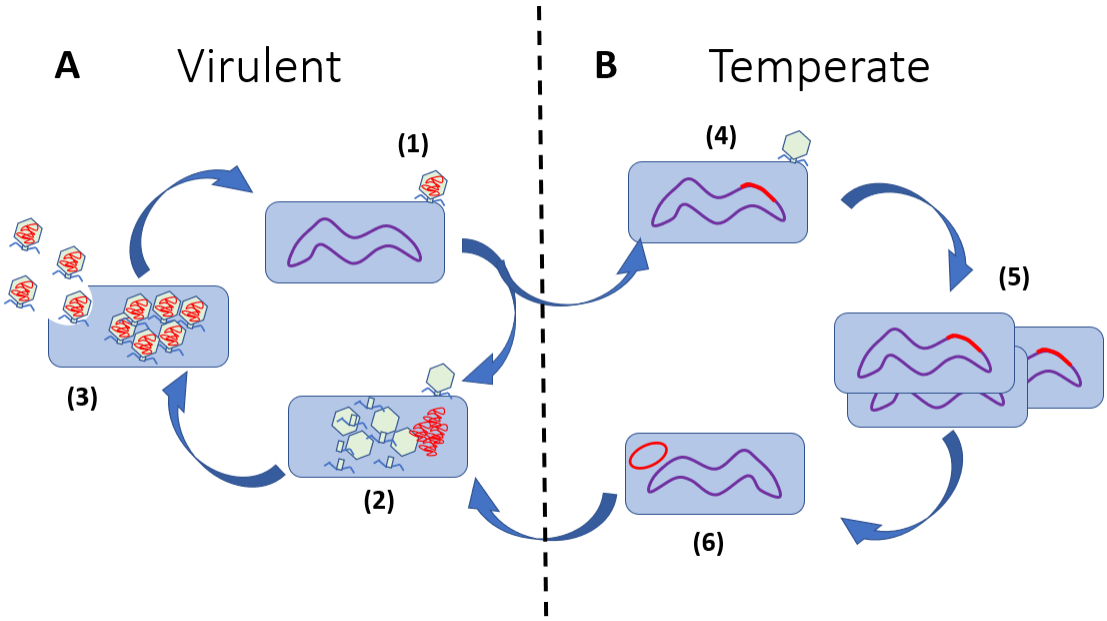


Figure 2 Virulent vs Temperate Phage Lifecycles
Illustration of the basic lytic (left) and lysogenic cycles (right). A) Virulent phage are restricted to the lytic cycle, steps (1-3). Step 1, phage recognizes and adsorbs to a specific receptor on the surface of its host cell. Step 2, phage DNA is transferred into the host cytoplasm and host takeover begins. Phage DNA and proteins are replicated and translated inside the infected cell. Step 3, virions are assembled and host cell lysis releasing new phage progeny for the cycle to continue. B) Temperate phage are capable of both the lytic cycle (1-3) and the lysogenic cycle (4-6). Lysogenic cycle step 4, after phage DNA is transferred into the host cell it is integrated into the genome of the host.

Step 5, phage DNA integrated into the host is repressed and the phage divides passively with the host. Step 6, a phage specific induction signal occurs, often host SOS response, causing phage DNA to excise from the host and enter the lytic cycle, step 2.

What Constitutes a "Healthy" Virome?

Relative to the bacteriome, very little attention has been paid to the viral component of the intestinal microbiome, or the virome. The virome is often barely discussed, if at all, in studies examining the intestinal "microbiome" despite the likelihood that it plays a key role in shaping bacterial composition of the gut microbiome (Andrey N Shkoporov & Hill, 2019). The lack of focus on the virome can be attributed to the lack of universally conserved features for taxonomic placement, as is possible with the 16S sequencing of bacteria, making it more difficult to group and quantify viral sequencing reads. For example, many meta-virome studies report ~70-90% of sequencing reads/contigs cannot be placed into functional or taxonomic categories, given the existing databases (Aggarwala et al., 2017). Additional barriers for metagenomic sequencing of the virome are the mosaic nature of bacteriophages, lack of suitable reference genomes, and the high degree of genes of unknown function within phage genomes, the "viral dark matter" (Aggarwala et al., 2017). As the virome is largely a function of the bacterial component of the microbiome given phage cannot persist without their cognate hosts, there exist a number of similar trends between the virome and the bacterial microbiome. Like the microbiome, the virome has been shown to be

variable between individuals, relatively stable over time, and broadly a reflection of an individual's microbiome (Samuel Minot et al., 2013; S. Minot et al., 2011; Moreno-Gallego et al., 2019; Reyes et al., 2010; A. N. Shkoporov et al., 2019). Multiple studies have shown the virome is remarkably stable, with some studies indicating the persistence of ~80% of assembled viral sequences for up to 2.5-years (the duration of the study) (Samuel Minot et al., 2013; A. N. Shkoporov et al., 2019). While there is consensus regarding the virome being highly specific to each individual, there is some controversy regarding the existence of a "healthy" or small "core" virome (Gregory et al., 2020; Manrique et al., 2016; Moreno-Gallego et al., 2019; A. N. Shkoporov et al., 2019). Some studies support of the existence of a "core" virome and report that an extremely small fraction, ~0.2-0.5% of assembled contigs or ~20 higher order taxonomic clusters, could be found in more than ~50% of the cohort participants (Garmaeva et al., 2021; Manrique et al., 2016; Moreno-Gallego et al., 2019; A. N. Shkoporov et al., 2019). However, the identification of a small core virome has been challenged (Gregory et al., 2020); therefore, this concept of a core virome requires additional study involving larger and more diverse cohorts. The majority of assembled virome contigs are highly variable between individuals, forming what Shkoporov et al. (2019) referred to as the persistent personal virome, or PPV, which constitutes dominant temporally stable phages commonly predicted to infect abundant bacteria in the gut. This PPV is thought to contribute to the inter-personal virome diversity between subjects (A. N. Shkoporov et al., 2019). Some studies have also indicated that the virome is dominated by virulent crAss-like phages of the gut bacteria *Bacteroides*, with crAss-like sequences accounting

for up to 90% of reads in some cases (Andrey N. Shkoporov et al., 2018a). This frequent virome domination by crAss-like phages is supported by numerous metagenomic virome studies (Dutilh et al., 2014; Guerin et al., 2018; Manrique et al., 2016; Moreno-Gallego et al., 2019; A. N. Shkoporov et al., 2019). In addition to the abundant crAss-like phages, these studies have also reported the order *Caudovirales* and *Microviridae* as the other predominant phage groups (Manrique et al., 2016; Samuel Minot et al., 2013; Moreno-Gallego et al., 2019; A. N. Shkoporov et al., 2019). However, it has been suggested that abundance of *Microviridae* may be somewhat exaggerated as an artifact of DNA amplification bias during the sample preparation (A. N. Shkoporov et al., 2019). However, a recent study reported the predominant members of the virome are *Siphoviridae*, *Myoviridae,* and *Podoviridae* members, rather than crAss-like phages and *Microviridae* (Garmaeva et al., 2021). The *Siphoviridae, Myoviridae,* and *Podoviridae* being large families of the order *Caudovirales* primarily distinguished by their morphology; however, their taxonomic utility is waning with an increase in the number of sequenced phage genomes, as it is clear they do not represent monophyletic clade. As such, these families will likely be dissolved in the near future for a more accurate taxonomic designation (Turner, Kropinski, & Adriaenssens, 2021). Nevertheless, a complete consensus has not yet been reached in the literature on which specific viral families are most dominant in the virome. Furthermore, like the presence of a core virome, there is also controversy as to which phage lifestyle predominates the virome, be it virulent or temperate. Early examinations of the virome have suggested it is dominated by temperate phages, as indicated by presence of integrases within assemblies (Reyes et

al., 2010). However, more recent studies suggest a virome dominated by virulent phages (Clooney et al., 2019; A. N. Shkoporov et al., 2019). Also like the microbiome, age-dependent shifts in the virome composition have been reported (Gregory et al., 2020; Efrem S. Lim et al., 2015). A study of infant twins through two years of age showed bacteriophage richness and diversity was greatest in newborns, primarily composed of the order *Caudovirales* and *Microviridae*. Diversity and richness decreased with age and the composition shifted to a decrease in *Caudovirales* and an increase in the relative abundance of *Microviridae* (Efrem S. Lim et al., 2015). Notably, crAss-like phages were only found in a single infant fecal specimen, implying colonization by the allegedly ubiquitous crAss-like phages happens later in life (Efrem S. Lim et al., 2015). These results are supported by a metanalysis of publicly available datasets ranging from 0 – 65+ year old individuals. This study also found a trend of crAss-like phage abundance increasing with age and an initially high abundance of *Caudovirales* declining into adolescence before increasing in abundance in adulthood (Gregory et al., 2020).

In summary, the individual virome is highly specific to each individual with each individual harboring on average >2,000 viral genomes (Garmaeva et al., 2021). However, our understanding of the virome is still in its infancy and many discrepancies exist in the literature; these discrepancies are likely due to fundamental differences in sample preparation and database-dependent vs independent forms of analysis. A summary of the general characteristics of the healthy virome can be seen in **Figure 1, right.** Additional cohorts, standardized methods, and increased capability for the taxonomic placement of the “viral dark matter” are necessary to obtain a clearer picture

on the “healthy” virome. Studies continue to identify novel intestinal phages from metagenomic sequences. Most recently Nayfach et al. (2021) assembled and deposited a staggering 189,689 genomes culling sequencing reads from over 11,000 human stool metagenomes. Consistent with the dominant members of the bacteriome, >75% of these genomes are predicted to infect Bacteroidia and Clostridia classes (S. Nayfach et al., 2021). Such mass phage assemblies will assist in the taxonomic placement and grouping of future virome metagenomic studies. Thus, the continuous refinement of methods and increases in sequencing capabilities will continue to illuminate the unknown aspect of the virome. However, despite our current fairly limited understanding of the “healthy” virome of the gut, a limited number of studies have also begun to assess the virome in its relation to a number of intestinal disease states.

Disease Associated Alterations to the Virome

While the impact of the bacteriome on various disease states has been studied in great detail with the aid of metagenomic and 16S sequencing, less investigation has been done to explore the virome in enteropathies associated with dysbiosis. Additionally, with bacteriophage, it is even more unclear if they contribute to the diseased state or are merely a consequence, given they require their cognate hosts presence in the microbiome to propagate. Nevertheless, a number of studies have attempted to catalog the alterations to the virome during diseased states, a summary of these findings can be seen in **Figure 1, right**. These alterations to the virome will be described here in brief, with additional focus given to IBD and EED as they pertain to ensuing chapters of this dissertation.

The virome of patients with CRC has been investigated to a limited degree with conflicting results in some instances. Hannigan et al. reported no significant changes in virome diversity (Hannigan et al., 2018), while Nakatsu et al. reported a significant increase in bacteriophage diversity amongst CRC patients (Nakatsu et al., 2018). Despite these conflicting reports on alterations to virome diversity, a number of differential operational viral units (OVU) were identified by Hannigan et al. (2018) and through random forest modeling a CRC virome signature could be established to predict the CRC disease state with similar accuracy to a bacterial microbiome-based model of CRC. However, additional study of the CRC virome is necessary to further validate any diagnostic potential of the virome as an indicator of disease. Examination of the virome in patients with recurrent CDI displayed significant increases in abundance of *Caudovirales* (Draper et al., 2018; Zuo et al., 2018), but with reduced diversity, richness, and evenness (Zuo et al., 2018). Additionally, Draper et al. reported a significant reduction of the CrAss-like phages in recurrent CDIs compared to healthy controls (Draper et al., 2018). In support of a role for phage in maintaining microbiome homeostasis, the viromes of recipients receiving FMT from healthy donors experienced alterations of the virome mirroring that of the donor with detectable transfer and recovery of a number of phages notably *Siphoviridae*, CrAss-like phages, and *Microviridae* from the donor for up to 12 months in some studies (Chehoud et al.; Draper et al., 2018; Zuo et al., 2018). After FMT, recipients also experienced a decrease in *Caudovirales* abundance and an increase in *Microviridae* (Zuo et al., 2018). In further support of a specific role of phage in resolving recurrent CDI, similar results to the FMT

were achieved by use of sterile fecal filtrate transfers (FFT), presumably only containing phage, small molecules, and metabolites from the donor (Ott et al., 2017). Thus, the success of FMTs in the treatment of recurrent CDI is likely due to, at least in part, of transfer of the donor virome. In the case of IBD, early examinations of the virome of IBD patients report a significant expansion in the richness of *Caudovirales* (Norman et al., 2015); subsequent studies have both supported (Fernandes et al., 2019) and contradicted these initial observations (Clooney et al., 2019; Zuo et al., 2019), leaving alterations to *Caudovirales* richness of the IBD virome in question. Zuo et al. did nonetheless report the UC mucosal virome was associated with an increased abundance of *Escherichia* and *Enterobacteriaceae* phages (Zuo et al., 2019). In addition to examining diversity, richness, and abundance, recent work by Clooney et al. also examined phage lifestyle and their relation to IBD. Key findings from this work reported rather than an alteration of virome richness, IBD patients display a loss of the virulent "core" virome, and an increase in temperate phages. This loss of virulent "core" results in healthy controls which are more similar to one another than to either the viromes of CD or UC patients, suggesting a dysbiosis of the virome in IBD (Clooney et al., 2019). This increase in temperate phages is exemplified by ~80% of differential viral clusters in IBD patients being predicted to be temperate, while only ~30% of the differential clusters from the healthy control cohort were predicted temperate (Clooney et al., 2019). Finally, it has also been suggested that independent of phage replication and host lysis, phages themselves may exacerbate inflammation in a TLR9-dependent manner leading to increased clinical signs of colitis with increases in IFN-y and IL-17A (Gogokhia et al.,

2019). However, while significant alterations were reported in this work, they were numerically minimal and the study used a combination of germ free mice as well as specific pathogen free mice, and extrapolating “normal” immune responses from naive immune systems can be challenging. In EED, analysis of the virome in children experiencing growth stunting and EED is currently very limited and requiring additional study to fully comprehend the status of the virome in the EED condition. However, the single study that does exist for EED reported no statistical difference in eukaryotic viruses, no significant changes in bacteriophage richness, and only a single significant change in phage diversity was found during the tested time period between the different growth velocities of children tested (Desai et al., 2020). Three *Caudovirales* phages were identified to be of differential abundance in the three different growth velocities tested: a siphoviridae phage infecting *Croceibacter*, an unclassified *Autographivirinae* phage, and an *Enterobacter*-infecting podophage (Desai et al., 2020). However, biological significance, if any, of these phages is unclear. While changes to the bacteriome and virome of EED were relatively minimal in this study, their interactions may be of clinical relevance and reflect dysbiosis associated with EED. Phage richness of the Adequate and Moderate growth velocity cohorts had a positive significant correlation with bacterial richness, while the phage richness of the poor growth velocity cohort had a negative correlation with bacterial richness (though not significant) (Desai et al., 2020), implying a disruption of phage and bacterial equilibrium in children exhibiting poor growth velocity.

In summary, there are clear alterations to the virome of diseases in which dysbiosis is a presumed contributor. It is unclear whether these alterations are merely a reflection of an altered bacteriome state, given phages require their cognate host for expansion, or if alterations to the virome through induction or exogenous acquisition of new phages are initiators to bacteriome disruption. However, given the complex environment of the intestine, multiple routes to bacterial phage resistance, and phage host range often restricted to strain levels, it is difficult to imagine how transiently acquired exogenous virulent phage at low levels could result in drastic reductions of key bacterial species. An additional obstacle to a better understanding of the impact of the virome is the limited resolution that can be achieved through metagenomic sequencing given the vastness of uncultured, unannotated, and uncharacterized phage; thus, database independent methods based on protein homology, as it seems is the current trend, are increasingly important to demystify the proportion of "unclassified" phages in which a large proportion of sequences can end up in. Such methods should also allow us to delve deeper past order level *Caudovirales* and family level *Microviridae* taxonomic classifications which is often the limit of analysis in prior studies.

Phage as Modulators of Microbiome

It has been a long-standing tenet that bacteriophages are important drivers of bacterial evolution and shape bacterial communities given their ability to kill and/or alter aspects of the host phenotypes (Koskella & Brockhurst, 2014). The effectiveness of sterile fecal filtrate transfer (FFT) in resolving CDI infections as described above

supports their natural ability to modulate the microbiome (Ott et al., 2017). Given these natural modulation abilities, it has been suggested that phages may hold the key to targeted engineering of the microbiome. The broad-spectrum nature of antibiotics makes them ill-suited for such precision as their repeated use can increase levels of dysbiosis and increase risk for obesity, diabetes, allergies, infections, CDI, as well as the enrichment of antibiotic resistance genes, reviewed thoroughly by Langdon et al. (Langdon, Crook, & Dantas, 2016). Phages are uniquely suited for precision engineering, however until recently with the advancements in various -omics techniques this novel role for phage application has been largely hypothetical. Early work in this field by Reyes et al. utilized gnotobiotic mice with a small defined consortium of bacteria and tested their response to an unknown phage consortium prepared from fecal filtrates (Alejandro Reyes, Meng Wu, Nathan P. McNulty, Forest L. Rohwer, & Jeffrey I. Gordon, 2013). This experimental system allowed for tracking of phage effects on defined community of commensal bacteria, but this resulted in limited utility given the unknown nature of the phage consortia and relatively small sample size. Nevertheless, this study reported a transient effect on bacterial community structure due to the simulated "phage attack" (Alejandro Reyes et al., 2013). More recently, phages were utilized to reduce asymptomatic intestinal carriage of uropathogenic *E. coli* strains (UPEC) and their effect on the microbiome structure compared to antibiotics (Galtier et al., 2016). Galtier and colleagues (2016) found that a phage cocktail was effective at reducing UPEC, and relative to antibiotics, bacteriophage had a smaller impact on bacterial diversity. Minimal impact on the non-target microbiome due to phage treatment

was also supported in a recent patient case study involving a *Staphylococcus aureus* infection (Mu et al., 2021). Thus, evidence supports minimal off-target effects on the microbiome from application of phage, however in some instances the conditions used to evaluate these off-target effects are less than ideal (i.e. an unknown phage consortia (Alejandro Reyes et al., 2013) or case studies (Mu et al., 2021)). The most thorough and controlled study, using heavily characterized bacteria and phages, on the ability of phage to modulate the intestinal microbiome was recently performed by Hsu et al (Hsu et al., 2019). Effects on the microbiome, phage survival, and metabolite profiles were analyzed in depth. Major findings from this study were that phage hosts can persist, often at low levels, in the presence of their cognate phage after an initial reduction reaching a sort of equilibrium between sensitive and resistant host bacteria. Second, there are off target effects on non-target bacteria as consortia bacteria exist as networks of interaction in which species may rely on metabolites or nutrients produced by other bacteria, therefore removal of specific bacteria may result in promoting or repressing effects as niches are opened for other bacterial species or key metabolites are lost (Hsu et al., 2019). However, while this analysis is foundational to our understanding of global impacts phages may have on the microbiome, it is difficult to say if such interactions would hold true in the presence of a naturally colonized gut microbiome consisting of hundreds of bacterial species at minimum, making the separation of different repression/promotion effects of individual species much more difficult. Nevertheless, these results demonstrate that repressive and/or promoting effects of the targeted bacterial host on other surrounding bacteria may also need to be considered. As we continue to gain a deeper

understanding of these bacterial interaction networks, it may also be possible to modulate non-cognate host bacteria in which infecting phage options are limited. Thus, this foundational study provides experimental evidence of phage's ability to modulate the microbiome and, in turn, the metabolome. Pulling from the evidence presented in this review for phage modulation, **Figure 3** represents a simplified hypothetical scenario in which phage might be used to bring the dysbiotic microbiota of a host back into a healthy state. In brief, a dysbiotic gut with high levels of expansion of certain bacteria not associated with the healthy gut microbiota, purple cocci and pink rods in this case, are exposed to a dose of infecting phage. The resulting phage predation significantly reduces the host bacterium (but not to elimination), and downstream bacterial network effects allow expansion or reduction in other bacterial competitors modulating the microbiome as well as the metabolome. This modulation of the microbiome, as well as metabolome, back to a "healthy" state in turn leads to a reduction in inflammatory effects and symptoms associated with diseased state.

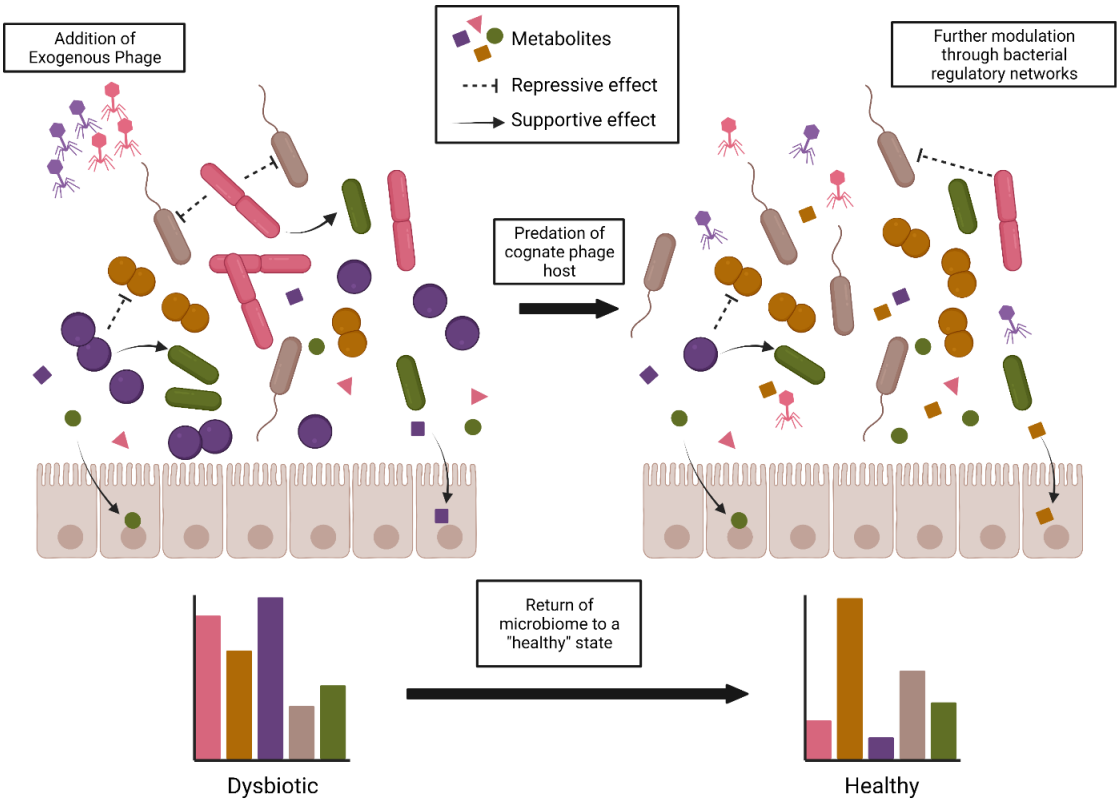


Figure 3 Model of Phage Modulation to Alleviate Microbiome Dysbiosis
Generally modified from findings of Hsu et al. (Hsu et al., 2019). Left, microbiome in a general dysbiotic state with arrows and dotted lines to indicate bacterial interaction repressing or supporting growth of other bacteria. Bacterial metabolome also reflects the abundance of bacterial drivers of dysbiosis. Below, relative bacterial taxa abundance corresponds color of bacterial taxa. Addition of phage targeting members of microbiome that drive dysbiosis, in this case pink and purple bacterial cells. Right, microbiome returned to a "healthier" state after knockdown of target bacteria. Reduction of target bacteria cause downstream effects further modulating the microbiome through bacterial interactions. Additionally, phage cause an alteration to the metabolome through the

modulation of the bacteriome. Return to a “healthy” state and presumably an alleviation of symptoms associated with the diseased state.

1.4. Conclusion and Continued Need for Individual Phage Isolation For Microbiome Modulation

While the multitude of metagenomic studies and controlled studies involving gnotobiotic mice are critical, they still do not eliminate the necessity of basic microbiology, cultivation, isolation, and individual study of various bacteria and bacteriophages. These metagenomic studies provide a detailed snap shot of the large trends in bacterial communities in various diseased states. These characterizations are key to identify bacterial targets and potential contributors to various diseased and healthy states. However, they do potentially miss minority members of the microbiome that may be below the detection limit, which may occupy a specialized niche and also contribute to dysbiosis linked diseased states. The question then becomes, now that we have characterized dysbiotic states, how best are we to act in order to bring the microbiome back to a healthy state of homeostasis? Isolation and culture of potentially impactful bacteria and their cognate phages will continue to be necessary to evaluate their basic biology and their function within the microbiome, tasks not fully achievable through only -omics methods. For example, the crAssphage, dominant in the intestinal microbiome, was initially identified via metagenomic sequencing (Dutilh et al., 2014), however it was not until years later through the cultivation and discovery of its cognate host that its unique biology has begun to be understood (Andrey N. Shkoporov et al.,

2018b; Andrey N. Shkoporov et al., 2021). Furthermore, the detailed study by Hsu et al. involving phage effects on host microbe interactions, was only possible due to their wise use of heavily studied bacteria and phage host pairs (Hsu et al., 2019). Bacteriophage represent a unique tool in the modulation of the gut microbiota; however, their isolation and characterization will continue to be necessary to minimize off target effects, avoid introduction of virulent genes, and maximize selection for phage with the highest potential for microbiome modulation. Furthermore, there is a great need for isolation of phages infecting difficult to cultivate members of the microbiota, as they are vastly understudied relative to phages infecting easily culturable model bacterial species. Each new phage isolated and characterized that infects a resident of the gut microbiota represents a new potential tool for modulating specific species of the microbiota with the intent of resolving dysbiosis. While there still remains a large knowledge gap in our understanding of phage-bacteria-host interactions, it may now be possible to move from mere observations of intestinal dysbiosis to targeted therapeutic interventions with the use of phage.

2. EFFECT OF CHRONIC AND ACUTE PATHOGENIC *E. COLI* CHALLENGE ON GROWTH PERFORMANCE, INTESTINAL INFLAMMATION, MICROBIOME, AND METABALOME OF WEANED PIGLETS

2.1. Introduction

Environmental enteric dysfunction (EED) is a chronic gut inflammatory condition in humans caused by continuous fecal-oral contamination of gastrointestinal pathogens. Chronic pathogen exposure results in condensed small intestine villi, increased crypt depth and inflammation of the lamina propria resulting in insufficient nutrient uptake (K. Watanabe & W. A. Petri, Jr., 2016). Poor sanitation procedures along with malnutrition are both positively associated with the development of EED in both children and adults (K. Watanabe & W. A. Petri, Jr., 2016). The gut microbiota of malnourished children tends to be reduced to an immature state, relative to healthy children, characterized by a reduction in diversity that is largely dominated by *Enterobacteriaceae* (S. Subramanian et al., 2014). Enterotoxigenic *Escherichia coli* (ETEC) are often associated with the development of EED (C. M. George et al., 2018). It is the most common source of diarrhea in young children under the age of five, and accounts for twenty to forty percent of traveler's diarrhea in adults (Qadri, Svennerholm, Faruque, & Sack, 2005). The onset of ETEC infection is marked by sudden watery stool which leads to dehydration as well as lethargy, lasting approximately three to four days (Qadri et al., 2005). Though the disease results in the reduction of gut microbiota

diversity, individuals affected by EED remain fairly asymptomatic (K. Watanabe & W. A. Petri, Jr., 2016).

A major physiological impact associated with EED is the stunting of growth, which may be caused by chronic nutritional deficiencies, persistent infections, or chronic inflammatory diseases such as EED (Owino et al., 2016). Stunting affected 171 million children under 5 years of age in 2010, and is predicted to affect 24% of young children in the developing world by 2020 (de Onis, Blössner, & Borghi, 2012). Early life stunting by enteric disease increases the likelihood of developing cardiovascular and metabolic diseases in adulthood (DeBoer et al., 2012). Furthermore, oral vaccines have reduced efficacy in EE patients and reduced neurocognitive development associated with chronic inflammation (K. Watanabe & W. A. Petri, Jr., 2016).

Therapeutic treatments against EED have not been highly successful. Antibiotic therapy and probiotic use have been investigated as possible treatment methods for children with EED; however, neither treatment improved intestinal barrier function compared to placebo (K. Watanabe & W. A. Petri, Jr., 2016). Therefore, there is increased interest to determine new methods to alter the populations of gastrointestinal pathogenic bacteria, and thus a suitable model is required to examine these intervention strategies.

Swine are used widely in biomedical research as a proxy for humans, due to similar physiology, immune systems, intestinal permeability, and intestinal enzymatic profiles. Weaned pigs represent a possible model for EED as porcine gastrointestinal anatomy, immune response, and ETEC clinical signs closely mimic that of humans

(Jiminez, Uwiera, Douglas Inglis, & Uwiera, 2015). Swine experiencing *E. coli*-induced post-weaning diarrhea (PWD) or inoculated with ETEC experience sloughing of intestinal villi, increased crypt depths, and scours (Fairbrother & Nadeau, 2019; Fairbrother, Nadeau, & Gyles, 2007; Isaacson, 1998; G. W. Jones & Rutter, 1972; Rhouma, Fairbrother, Beaudry, & Letellier, 2017). It has been shown that ETEC infections in weanling pigs can be caused by a single dose of approximately 10^9 CFU (Berberov et al., 2004; Wellock, Fortomaris, Houdijk, & Kyriazakis, 2008). However, this high acute single dose most likely does not accurately represent the real-world scenario of EED in humans or PWD in piglets in which initial infection is caused by chronic exposure to lower doses of *E. coli* as can be found in contaminated feed, water, soil, and elsewhere in the environment (Fairbrother & Nadeau, 2019). The objective of the present study was to develop a pathogenic *E. coli* induced EED weaned swine model that can be used to evaluate different EED therapies. Given the chronic and often asymptomatic nature of EED, both one-time acute and chronic dosage models were evaluated. To our knowledge, this is the first reported comparison on the effects of a one-time high acute dose vs a lower chronic daily dose in an animal model. Comparing the single- or repeated-dose models for studying chronic inflammatory gastrointestinal disorders is critical to being able to successfully evaluate potential treatments for use in humans. The *E. coli* strain used in this work encodes F18ac+ fimbrial adhesin, heat-labile enterotoxin IIA (LT-IIA), heat-stable enterotoxin II (STIIB), as well Shiga toxin (Stx2e), technically identifying it as both ETEC and STEC. However, there was no evidence of the edemas classically associated with the presence of Stx2e, making

significant expression of Shiga toxin under the trial circumstances questionable; thus, the trial strain will simply be referred to as a pathogenic *E. coli* strain. Ultimately, the trial conditions presented here did not mimic in full the clinical signs of EED; for that reason, the model presented here may more accurately depict a general *E. coli*-induced enteropathy. The detailed model presented here of *E. coli* induced gastrointestinal disease provides valuable insights for future novel treatments for human enteropathies and post-weaning diarrhea common amongst piglets.

2.2. Results and Discussion

Growth performance

Pigs used as an experimental model for PWD were challenged with a spontaneous nalidixic acid-resistant mutant of *Escherichia coli* strain NCDC 62-57 (ATCC 23545) referred to hereafter simply as *E. coli* 62-57nal in either a single acute high-titer dose (~10^9 CFU), or in a series of daily lower-dose challenges (~10^7-10^8 CFU). This strain was chosen for its ability to cause illness during preliminary trials amongst F18+ and K88 (F4) ETEC strains. All pigs were held for two days prior to the start of the trial and were asymptomatic for gastroenteritis. Additionally, pigs were not colonized by organisms capable of forming colonies on MacConkey agar amended with 50 µg/ml nalidixic acid (MacConkey+nal), and no endogenous phage infecting *E. coli* 62-57nal were identified. Thirty-six presumptive coliform colonies from pooled fecal samples plated on MacConkey agar (0 µg/ml Nal; three colonies per pen) were also tested by PCR for the presence of genes encoding Shiga toxin type 1 (Stx1), Shiga toxin type 2 (Stx2), heat-stable enterotoxin I (ST1), heat-stable enterotoxin II (ST2) and heat-labile

toxin (LTI). All colonies were negative for genes encoding Stx1, Stx2, ST2 and LTI, but three colonies were positive for ST1. Presence of the ST1 gene, *estA*, alone is not a strong predictor of ability to cause disease (Berberov et al., 2004; Wassenaar & Gunzer, 2015; W. Zhang et al., 2006) and pigs were asymptotic, so all animals were retained in the study.

In general, pigs administered *E. coli* 62-57nal via both the acute and chronic dosing regimens presented similar clinical signs with the majority of pens developing scours by day 2 and continuing through day 6. Control pens had visibly soft feces on day 5 and 6 with a single incident of scours on day 6, however the animals in control pens remained visibly healthy throughout the trial period. Additionally, the control pen with the incidence of scours was culture negative for the inoculated *E. coli* 62-57nal throughout the trial, so scours may have been induced by stress or other native microbiota. There was no evidence of difference for overall average daily gain (ADG), average daily feed intake (ADFI), and gain:feed (G:F) of the different treatment groups ($P = 0.184$**,** Table 1). However, there were numerical differences between pigs fed the treatments, suggesting that the modest number of replicates and the inherently high post weaning variability in performance was responsible for the failure to detect significant differences in growth performance between treatments/pens. This lack of evidence for significant growth differences is similar to previously reported results (Bhandari, Xu, Nyachoti, Giesting, & Krause, 2008). Pigs administered the acute and chronic dose of *E. coli* 62-57nal had a 54.7% and 14.9% reduction in ADG compared to the control pigs, respectively (Table 1). The control group had the lowest ADFI among the three

treatments with acute and chronic dosing regimens increasing feed intake by 17.3% and 29.95%. These findings are in agreement with previous work that showed a 24% decrease in control pigs ADFI compared to the pigs inoculated with ETEC O149 on d 3 to d 6 (Wellock et al., 2008). Madec et al. had similar results with a decrease in weight of weaned piglets inoculated with pathogenic *E. coli* expressing K88 fimbriae from day 0 to day 2 which then recovered by day 9 of the trial (Madec et al., 2000). In this study, the acute challenge group had the poorest mean G:F conversion with the control group having the highest mean feed efficiency. Piglets experiencing PWD have been reported to exhibit reduced weight gains (Casewell, Friis, Marco, McMullin, & Phillips, 2003; Fairbrother et al., 2007). Nevertheless, statistically significant reductions in weight performance were not observed, perhaps due to the relatively brief duration of the trial or small sample sizes.

Table 1. Effects of *E. coli* 62-57nal treatment on nursery pig performance

	Treatment				
Item	**Control**	**Acute**	**Chronic**	**SEM**	**Probability, (*P*)**
d 0 to 6					
ADG, g	263	119	225	55.93	0.224
ADFI, g	179	210	232	43.05	0.688
G:F	2.02	0.66	0.89	0.722	0.184

Table 1. Growth Performance.

A total of 24 weaned barrows (initial BW approximately 14 lb; Landrace × Large White) were used in a 7 d study. Pigs were housed 2 pigs per pen with a total of 4 pens per treatment. Treatments refer to the following: Control= Pigs administered a sham challenge of PBS each day, Acute = pigs administered approximately 10^9 CFU of *E. coli*

62-57nal on d 1 then sham challenged for the remaining time, Chronic = pigs administered approximately 10^7 CFU of *E. coli* 62-57nal from d 1 to d 6.

Table 2. Diet Composition (as-fed basis). Composition of diet fed to pigs ad libitum, represents a standard phase 1 nursery pig pelleted diet. Formulated to meet or exceed National Research Council standards (Council, 2012).	
Table 2. Diet composition (as-fed basis)	
Ingredient	**%**
Corn	36.65
Soybean Meal (46.5.% CP)	10.6
Fermented soybean meal	10
Milk, Whey Powder	30
Fish Meal Combined	10
Soybean Oil	1
Limestone	0.5
Sodium chloride	0.3
L-Lys-HCL	0.265
DL-Met	0.175
L-Thr	0.105
Trace mineral premix[1]	0.15
Vitamin premix[2]	0.25
Phytase[3]	0.02
Total	100
[1]Provided per kilogram of diet; 110 mg of iron, 110 mg of zinc, 33 mg of manganese, 17 mg of copper, 0.3 mg of iodine, and 0.3 mg of selenium (Kansas State University)	
[2] Provided per kilogram of diet; 1,102 IU of vitamin A, 1,378 IU of vitamin D_3, 44 IU of vitamin E, .04 mg of B_{12}, 4.4 mg of menadione, 8.3 mg of riboflavin, 28 mg of pantothenic acid, and 49 mg of Niacin (Kansas State University).	
[3] Ronozyme HiPhos (DSM Nutritional Products, Parsippany, NJ) 540 FYT/kg feed, with a release of 0.13% available P.	

Table 2. Diet Composition.

Bacterial colonization and fecal shedding

The ability of *E. coli* 62-57nal to colonize the gastrointestinal tract of inoculated piglets was determined by measuring colony-forming units recovered from intestinal mucosa, intestinal luminal contents, and in feces. Inoculated strain counts adherent to the mucosal lining were found to be variable, with ~50% of samples, ranging from 0.21 to 1.71 g of intestinal scraping, yielding counts above the detection limit (5,000 CFU/ml of tissue homogenate). Of the samples yielding enumerable colonies, bacterial counts ranged from 1.4×10^3 to 4.4×10^7 CFU/g in the duodenum, jejunum, ileum, cecum and colon (**Figure 4A**). Bacterial counts in the cecal and colonic luminal contents, ranging from 0.14 to 11.07 g of digesta, were more reliably above the detection limit and ranged from 1.2×10^3 to 7.2×10^8 CFU/g, suggesting bacterial proliferation in the unattached population.

Acute and chronic treatments had higher prevalence of *E. coli* 62-57-nal in feces (~10^5 to 10^7 CFU/g) than control pigs on all sampling days (**Figure 4B**). Control pen animals sporadically shed *E. coli* 62-57-nal in the feces at levels near the lower detection limit (500 CFU/g), likely reflecting low levels of pen cross-contamination. Pigs administered the acute pathogenic *E. coli* dose exhibited significantly higher fecal shedding on day 1 (~10^7 CFU/g, $P = 0.001$) compared to the chronic dose, though there was no statistically significant difference in fecal *E. coli* 62-57nal counts between the acute and chronic treatments after d 1. This result is consistent with other piglet studies which observed peak shedding between 24 and 48 h post-inoculation (P. H. Jones, Roe, & Miller, 2001; Owusu-Asiedu, Nyachoti, & Marquardt, 2003). Pathogen shedding in the acute group

remained high through d 6, indicating that the *E. coli* 62-57nal was able to persistently colonize the gastrointestinal tract of swine.

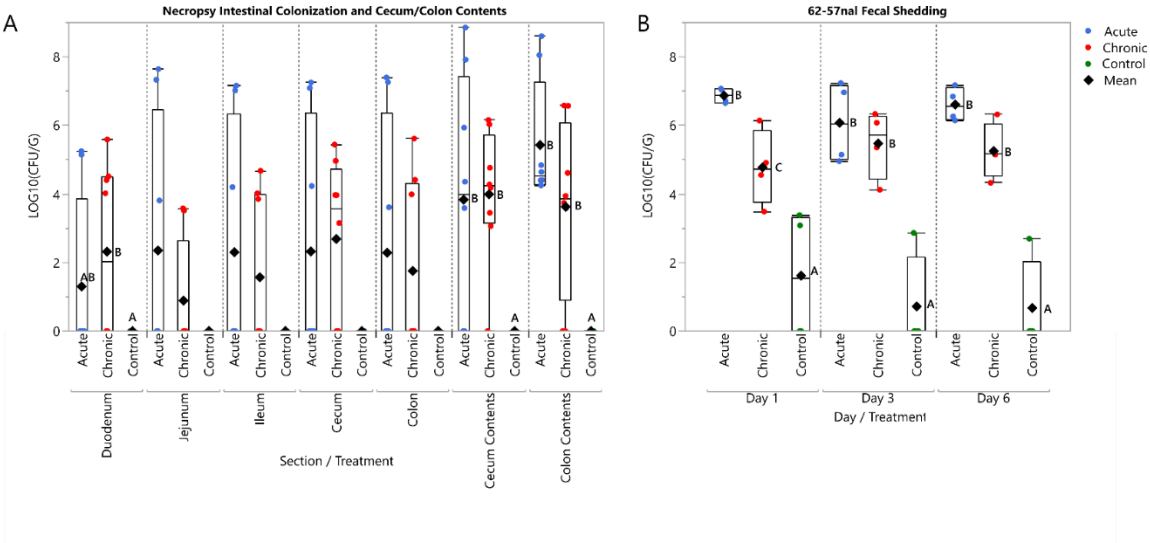


Figure 4. Acute and Chronic Regimens Are Sufficient to Cause Colonization of *E. coli* 62-57nal.
(A.) At time of necropsy, intestinal scrapings from the duodenum, jejunum, ileum, cecum, and colon were collected and plated for bacterial enumeration. Additionally, cecum and colon contents were collected for direct plating. (**B.)** Fecal samples were collected from each pen on day 1, 3, and 6. These samples were serial diluted and plated for direct enumeration of bacteria. LOD~$5x10^2$ CFU/g. **Both (A.) and (B.)** Circles indicate sampling data points, diamonds indicate mean. Sample data was transformed using LOG10. Treatment means grouped by day/section with different letters were significantly different ($P<.05$); those lacking letters indicate no level of significance.

Markers of inflammation and intestinal leakage

Infection-induced inflammation is mediated by increased levels of pro-inflammatory cytokines (Arango Duque & Descoteaux, 2014). To evaluate the level of *E. coli* induced inflammation, serum interleukins 6 and 8 (IL-6 and IL-8) were measured. IL-6 and IL-8 are useful biomarkers since they have been linked to intestinal inflammation (Gabay, 2006; Grimm, Elsbury, Pavli, & Doe, 1996; Guerrant, DeBoer, Moore, Scharf, & Lima, 2013; Scheller, Chalaris, Schmidt-Arras, & Rose-John, 2011) and have been employed as markers of porcine immune activation in response to *E. coli* previously (Fossum, Wattrang, Fuxler, Thorleif Jensen, & Wallgren, 1998; C. Y. Lee, Kim, Park, & Han, 2017; McLamb, Gibson, Overman, Stahl, & Moeser, 2013; Nyachoti, Kiarie, Bhandari, Zhang, & Krause, 2012). On d 6 of the study, pigs challenged with *E. coli* 62-57nal had increased ($P < 0.05$) concentrations of serum IL-6 compared to control pigs (**Figure 5**). However, there was no difference in IL-6 concentrations between acute and chronic treatments. Similar elevations of IL-6 were also observed in the treatment groups of a bacterially induced murine model of chronic intestinal inflammation (E. M. Brown et al., 2015). For concentrations of IL-8, there was a marginally significant overall treatment effect on d 6 ($P = 0.089$); chronic pigs had increased ($P < 0.05$) serum IL-8 concentrations compared to control pigs, and acute dose pigs were intermediate ($P =0.5423$). Lee et al. (2017) observed peak serum IL-8 levels in ETEC-challenged piglets between day 0 and 2 which then declined through d 7. This could explain the lower levels of serum IL-8 in the acute challenge group as serum cytokines were measured 6 d after acute challenge, while the daily chronic challenge may maintain elevated IL-8 concentrations. Increased levels of IL-6 and IL-8 in response to challenge with *E. coli* 62-57nal is in agreement with prior

work demonstrating these cytokines as markers of inflammation and infection (E. M. Brown et al., 2015; C. Y. Lee et al., 2017).

Increased concentrations of fecal calprotectin have been positively correlated with the histological activity of inflammatory bowel disease in humans (Roseth, Aadland, Jahnsen, & Raknerud, 1997), and serve as a marker of neutrophilic intestinal inflammation (Costa et al., 2003). Fecal calprotectin has also been suggested as a noninvasive marker of intestinal inflammation in swine (Barbosa et al., 2021), though additional study is necessary. Past studies have investigated calprotectin in swine plasma, intestinal lumen, and jejunal mucosa, all of which were found to be positively correlated with bacterial infection (Splichal, Fagerhol, Trebichavsky, Splichalova, & Schulze, 2005; Xiao et al., 2014). In the present study, there were no significant treatment effects on fecal calprotectin concentration. This is the first study to our knowledge to test fecal calprotectin in pigs challenged with *E. coli*, and this indicates fecal calprotectin may not be an informative biomarker in this model.

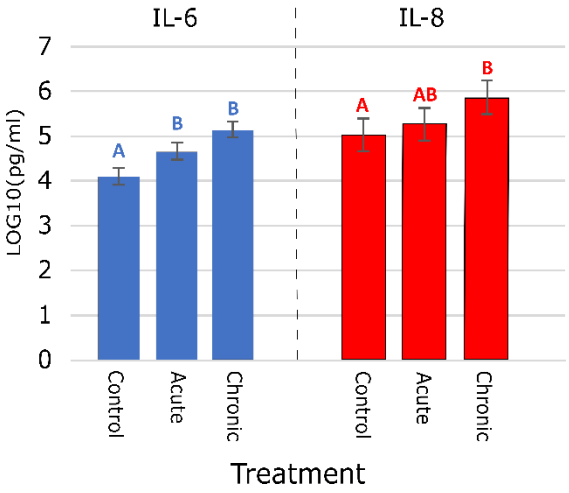


Figure 5. Effect on Serum Cytokines IL-8 and IL-6.
On day 6, serum samples were taken from each piglet. Samples were tested by ELISA and data was transformed using $\log_{10}$. Means with different letters within each cytokine differ significantly ($P < 0.05$). Error bars represent SEM.

Villi length and histopathology

Previous studies involving swine challenged with ETEC strains have reported villous atrophy and reductions in crypt depth (Gao et al., 2013); similar symptoms have also been reported in chronic intestinal enteropathies in humans (K. Watanabe & W. A. Petri, Jr., 2016). At time of necropsy, sections were collected to evaluate villi length in the piglet model, but no morphologic changes were observed between treatment groups. General bacterial rod attachment was evaluated by an anatomic pathologist and observed sporadically in all samples with no apparent correlation between rod attachment and direct bacterial plating as only 38% (11/29) of samples with rod attachment tested positive for *E. coli* 62-57nal by direct plating. Villus length in *E. coli* challenged animals did not differ from the controls in the duodenum (P=0.7125), jejunum (P=0.3719), and

ileum (P=0.778). Representative tissue sections can be seen in **Figure 6**. Lack of villus blunting may be due to the limited duration of this study. A prior longer-term study (21 d) with a murine model of chronic intestinal inflammation obtained villus blunting through a combination of bacterial challenge and malnutrition (E. M. Brown et al., 2015). Similarly, post-weaning anorexia in piglets has been shown to be associated with reduced villus heights (McCracken, Spurlock, Roos, Zuckermann, & Gaskins, 1999). Therefore, given a longer trial period and/or malnourishment, blunting may have been eventually observed in our present model. Histology is also only able to evaluate a tiny fraction of the intestinal tract, so lesions must be broadly distributed throughout the tissue to be detectable by this method. Based on this data, histologic analysis does not appear to be a useful method for evaluating this model.

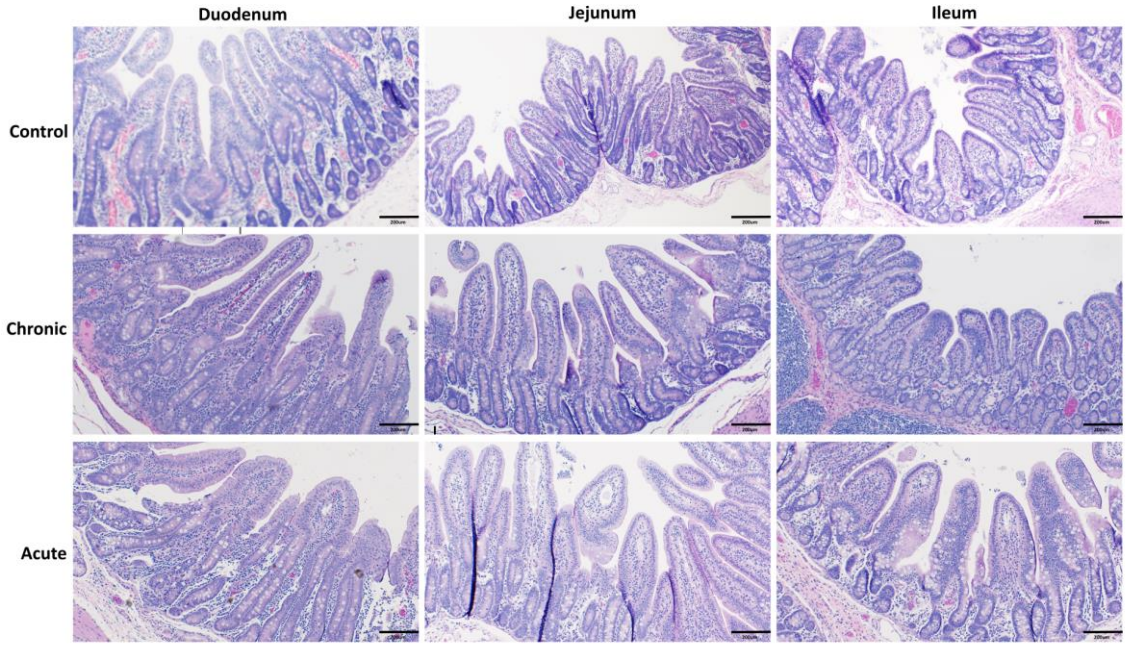


Figure 6. Representative Tissue Sections.
Representative Tissue Sections. from the duodenum, jejunum, and ileum from control, chronic, and acute treatment groups. Sections stained with hematoxylin and eosin. Scale bar = 200 μm.

Effects on the microbiome by 16S qPCR analysis

To observe any changes of the gut microbiota caused by our acute or chronic dosing treatments, targeted 16S qPCR was performed for select bacterial groups on fecal samples collected from pens at days -1, 1, 3, and 6. Relative abundances obtained were consistent with previous examinations of the piglet microbiome, showing a microbiome dominated by Bacteroidetes and Firmicutes (L. Chen et al., 2017; Dou et al., 2017). Overall, the bacterial groups tended to increase relative to control and pre-treatment samples, likely due to natural microbiome succession. A summary of these significant (*P*

< 0.05) or marginally significant ($P < 0.10$) bacterial group changes at each time point is shown in **Table 3**. Both acute and chronic doses of *E. coli* 62-57nal impacted relative quantities of *E. coli* populations compared to the control. Additionally, both dosing regimens had comparable impacts on microbiome progression. Pre-treatment compared to post-treatment samples of the acute dose had the most significant/marginally significant changes with eight of the ten tested bacterial taxa (Bacteroidetes, *Enterococcus*, *Faecalibacterium*, Firmicutes, *Lactobacillus*, *Streptococcus*, *Fusobacterium*, and Universal) showing increased populations. The chronic treatment showed similar but less dramatic changes, with six of ten taxa (Bacteroidetes, *Enterococcus*, *Lactobacillus*, *Streptococcus*, *E. coli*, and *Ruminococcaceae*) showing increased levels from pre- to post-treatment. The control group showed only two altered bacterial groups, *Enterococcus* and *E. coli*, from pre- to post-treatment. The observed increases for *Enterococcus* and *E. coli* within the control treatment are consistent with the previously reported natural post-weaning piglet microbiome maturation which shows an increase in levels of *Enterococcus* and *Enterobacteriaceae* at 8 d post-weaning (Montagne et al., 2007). The acute dose of *E. coli* had a slightly more pronounced impact on the gut microbiome maturation than the chronic dose, however both acute and chronic treatments were sufficient to cause a detectable dysbiosis.

Principal component analysis (PCA) of 16S qPCR results also provides clear evidence of dysbiosis in *E. coli*-treated groups (**Figure 7A**). PCA of pre- vs post-treatment samples indicates that by d 6 the acute treatment clearly clustered away from its pre-treatment sample, while the chronic day 6 sample showed an intermediate

clustering from its pre-treatment sample. In contrast, the control group remained tightly clustered throughout the trial period. This contrast in clustering suggests the microbiome perturbations are induced by the *E. coli* challenge and are not merely normal microbiota progression. The observed relative stability of the control microbiome is consistent with other studies, which reported a microbiome shift immediately after weaning and reached relative stability within 10 days after weaning (L. Chen et al., 2017). These findings indicate that both the single acute dose and the chronic lower dose of *E. coli* 62-57nal caused varying degrees of a similar dysbiosis. While in this present study the acute dose of pathogenic *E. coli* provided a more pronounced microbiome defect, the slight alteration caused by the chronic dose may still be more reflective of chronic exposure to pathogens, as piglets may experience.

Bacterial Groups that change from Control at Day 1		Bacterial Groups that change from Control at Day 3		Bacterial Groups that change from Control at Day 6		Bacterial Groups that change at Day 6 from their corresponding pretreatment sample at Day -1		
Acute	**Chronic**	**Acute**	**Chronic**	**Acute**	**Chronic**	**Acute**	**Chronic**	**Control**
Fusobacterium* ▼(.0571)	Fusobacterium* ▼(.0571)	E. coli** ▲(.0286)	Faecalibacterium** ▲(.0286)	Fusobacterium** ▲(.0286)		Bacteroidetes** ▲(.0286)	Bacteroidetes** ▲(.0143)	Enterococcus** ▲(.0286)
	Enterococcus* ▼ (.0571)	Streptococcus** ▲ (.0286)	Universal** ▲(.0286)	Ruminococcaceae** ▲(.0143)		Enterococcus** ▲(.0286)	Enterococcus** ▲(.0143)	E. coli* ▲(.0571)
	E. coli* ▲(.0571)		Firmicutes* ▲(.0571)	Faecalibacterium* ▲(.0571)		Faecalibacterium** ▲(.0286)	Lactobacillus** ▲(.0143)	
			Bacteroides* ▲(.0571)	E. coli* ▲(.0571)		Lactobacillus** ▲(.0286)	Streptococcus** ▲(.0286)	
			Lactobacillus* ▲(.0571)			Firmicutes** ▲(.0286)	E. coli* ▲(.0571)	
			E. coli* ▲(.0571)			Universal** ▲(.0286)	Ruminococcaceae* ▼(.0571)	
						Streptococcus* ▲(.0571)		
						Fusobacterium* ▲(.0571)		

Table 3. qPCR Summary Table.
LOGSQ treatment mean compared to Control at each time point (Left) and LOGSQ treatment mean at day 6 compared to its corresponding pretreatment sample at Day-1 (Right). Non-parametric Wilcoxon exact test performed on LOGSQ. Only those significant or marginally significant listed here, significant defined at $P<.05$ and marginally significant at $P<.10$. (▲) or (▼) indicate mean LOGSQ are greater than or less than relative to control or Pretreatment. Significant** Marginally Significant*

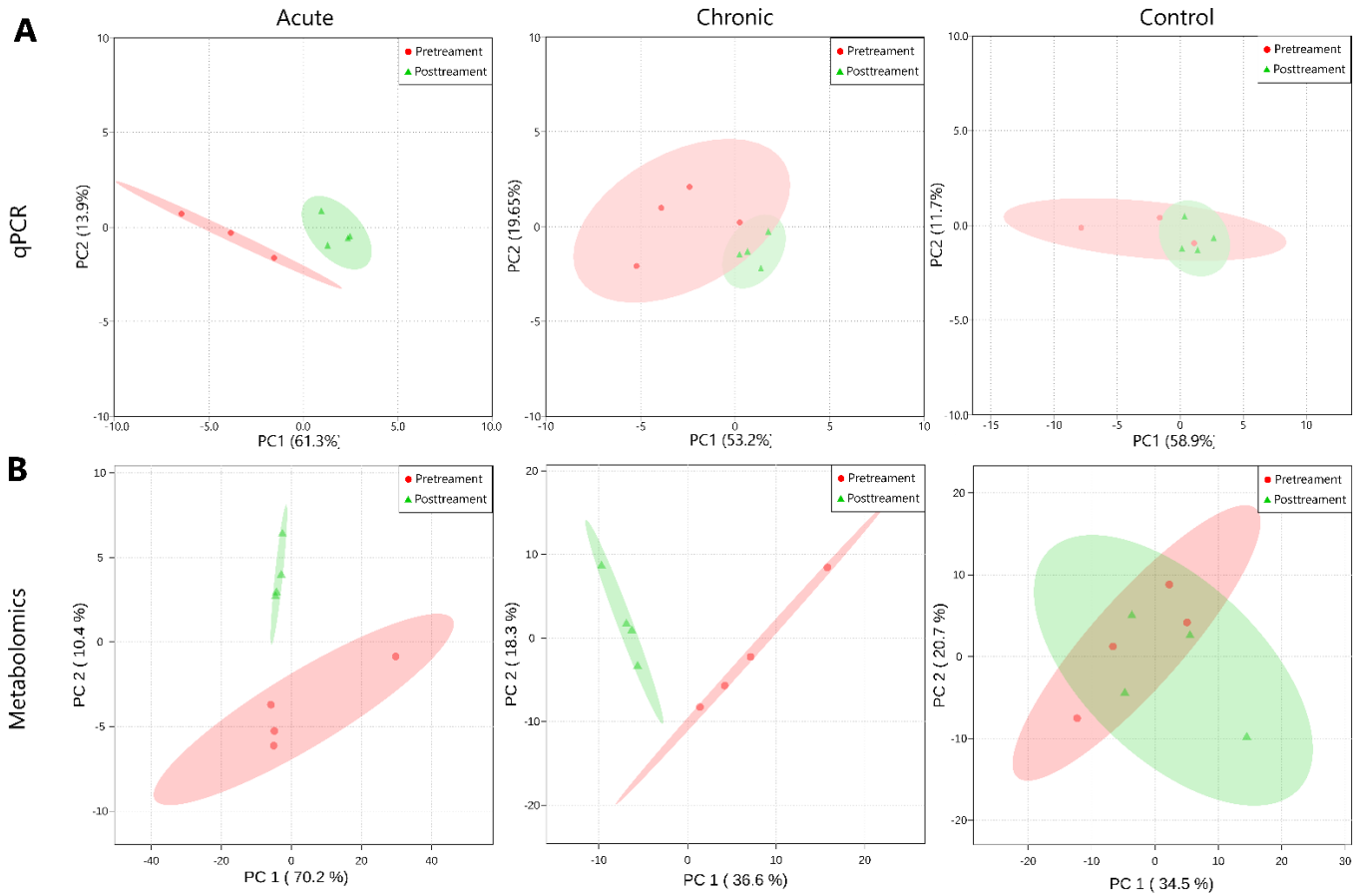


Figure 7. Clustering of Treatments by Principal Component Analysis (PCA) of Metagenomic and Metabolomic Results.

By the end of the trial period, both chronic and acute treatments are able to be separated from their respective pretreatment samples using qPCR and metabolomic data while the control group remains relatively constant. **(A.)** PCA comparing pre-treatment and post-treatment samples using 16S qPCR of major bacterial taxonomic groups. Red dots indicate pre-treatment samples, green diamonds indicate post-treatment samples, and green/red ellipses represent 95% confidence regions. **(B.)** PCA comparing pre-treatment and post-treatment samples using identifiable fecal metabolites.

Alterations in the fecal metabolome

To further characterize the differences in disease state caused by acute and chronic *E. coli* 62-57nal challenge, untargeted metabolomics was performed on fecal samples collected pretreatment (d -1) and d 1, 3 and 6 post-treatment. Metabolite profiles of fecal samples were analyzed by Metaboanalyst (Xia & Wishart, 2016). Analysis of the identifiable metabolites by PCA clearly distinguished between challenge and control groups (**Figure 7B**). Similar to the results of microbiome analysis (**Figure 7A**), the acute and chronic day 6 samples clearly cluster separately from their pre-treatment samples, while the control samples did not separate. The stability of the control group indicates the natural enzymatic, microbial, and structural maturation of the weaned piglet gut (Montagne et al., 2007) was not responsible for the observed shifts in the acute or chronic treatment groups.

As with the 16S qPCR approach, metabolomic comparison of post-treatment samples with their respective pre-treatment samples was more informative when identifying significant changes in individual metabolites. Volcano plots ($P<0.10$ and >2-fold change) were used to identify metabolites that significantly changed following treatment (Xia & Wishart, 2016) (**Figure 8**). A full list of metabolites identified by volcano plot is provided in **Table 4**. Metabolomics data related to swine, while growing, are fairly limited (Banerjee, Carmelo, & Kadarmideen, 2020; Carmelo, Banerjee, da Silva Diniz, & Kadarmideen, 2020; Chang et al., 2021; Gong et al., 2017; Y. Li et al., 2018), particularly metabolomics pertaining to piglets in a diseased stated, most of which are serum or tissue based (He, Liu, & Ji, 2021; Ren et al., 2015; Sugiharto,

Hedemann, & Lauridsen, 2014; Wu et al., 2018). The Human Metabolome Database (Wishart et al., 2018) is much more comprehensive, particularly for diseased states, than the Livestock Metabolome Database (Goldansaz et al., 2017). Given these limitations in existing swine metabolomics data, the metabolites identified in this manner were categorized based on the Human Metabolome Database chemical taxonomy and compared with relevant human diseased states in which fecal metabolomics are more prevalent. Only four identified metabolites were shared between the chronic and control groups, and ten were common to both the chronic and acute treatment groups (**Figure 8**). Changes in metabolites in the control group were presumed to be associated with the normal development of the weaned piglet gastrointestinal tract.

Within the chronic and acute treatments, increased levels of amino acid metabolites were identified in the post-treatment samples, including lysine, ornithine, homoserine, histidine, tyramine, beta-alanine, (**Figure 8, Table 4**). The increased levels of amino acids and amino acid metabolites in post-treatment fecal samples suggests the *E. coli* 62-57nal treatment led to amino acid malabsorption and/or secretion, likely due to disruption of chemical gradients, inflammation, and microbiome perturbations within the gut caused by pathogenic *E. coli* treatment. Additionally, metabolites associated with bacterial amino acid degradation, 5-aminovaleric acid and putrescine, were found at increased levels within acute and chronic treatment samples. These have been previously associated with ulcerative colitis (Fothergill & Guest, 1977; Le Gall et al., 2011). The presence of these metabolites is consistent with the model that suggests inflammation caused by *E. coli* treatment induced amino acid malabsorption and subsequent

degradation by the resident microbiota. Increased presence of fecal amino acids is supported by Wu et al. who reported decreased amino acid metabolites in jejunal tissue from ETEC-challenged piglets (Wu et al., 2018), supporting malabsorption in small intestine leading to increased presence of amino acids in the feces. This observation of increased fecal amino acid and amino acid metabolite levels also agrees with studies examining fecal metabolite profiles of humans with inflammatory bowel diseases like ulcerative colitis and Crohn's disease of which chronic diarrhea is a symptom (Bjerrum et al., 2015; Jansson et al., 2009).

Fecal metabolites that were significantly reduced in post-*E. coli* treatment samples were primarily fatty acid metabolites (**Figure 8**), including stearic acid, myristic acid, and arachidic acid. The levels of lipid-soluble alpha-tocopherol (vitamin E) was also reduced in both treatment groups. The observed depletion of fatty acids within the feces may be indicative of immune system activation, which is supported by our observation of increased levels of serum IL-6 and IL-8. Growing evidence on the importance of "immunometabolism" suggests activated macrophage subtypes and chronically activated T-cells demonstrate increased uptake of fatty acids as they rely more on fatty acid oxidation in order to maintain the high energy levels required to mount an immune response (Byersdorfer, 2014; Namgaladze & Brune, 2016). Prior studies examining metabolomic profiles of human inflammatory bowel diseases found pronounced decreases in the levels of short chain fatty acids (SCFA), which are the end products of bacterial fermentation that are absorbed by the large intestine; this presumably signaled a dysbiosis of gut flora (Bjerrum et al., 2015; Jansson et al., 2009).

In our current study a reduction in the SCFA metabolites butyrate, alpha-ketoglutarate and fumaric acid were observed. Our metabolomic findings indicate both the chronic and acute *E. coli* 62-57nal treatments caused sufficient dysbiosis to statistically distinguish pre- and post-treatment samples (**Figure 7B**) in large part due to amino acid malabsorption and reduction in fatty acid metabolites (**Figure 8**), generating metabolomic profiles resembling those of human inflammatory gastrointestinal disease associated with diarrhea.

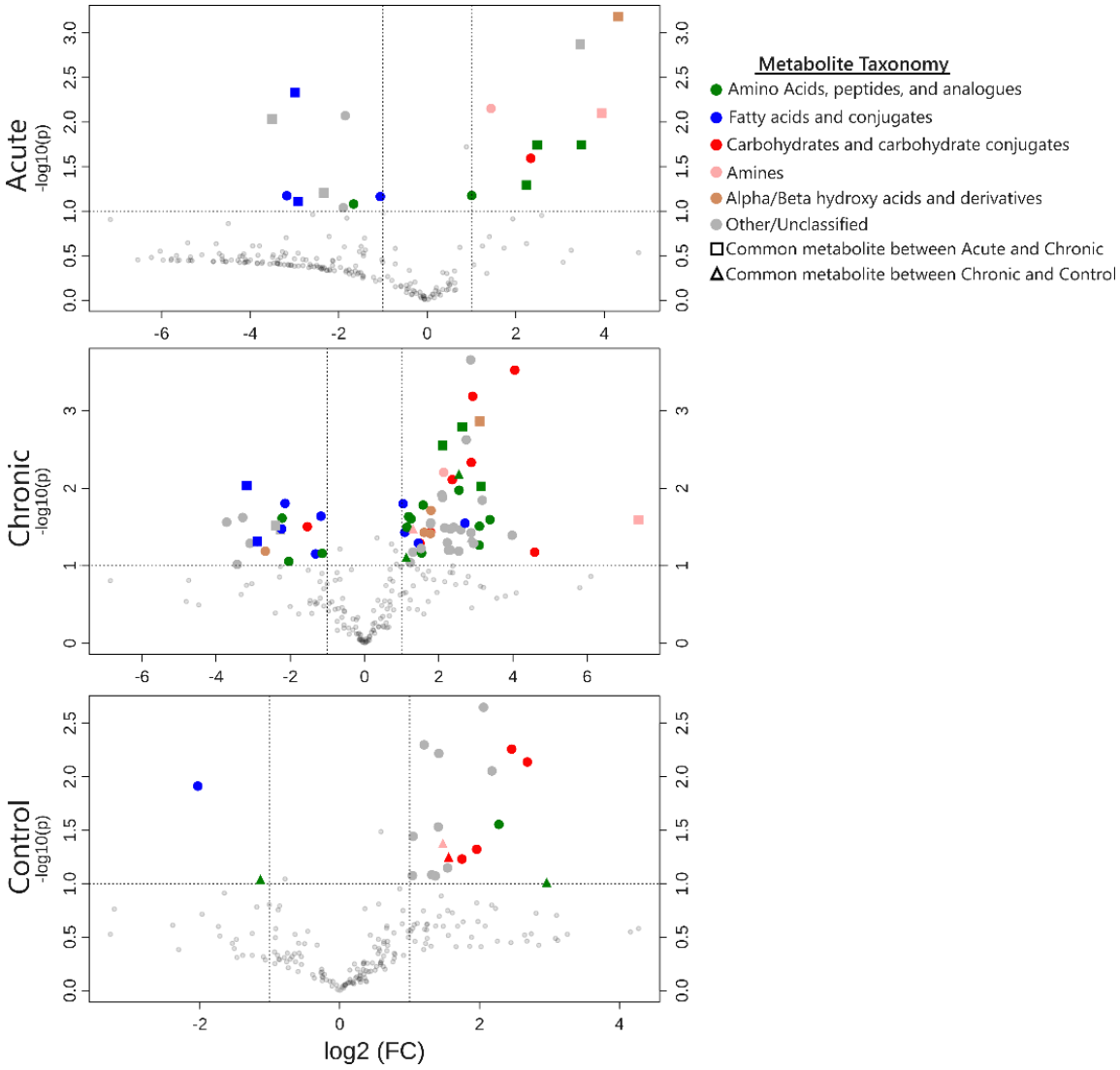


Figure 8. Volcano Plots Comparing Changes in Identifiable Metabolite Profiles in Pre- and Post- Treatment Fecal Samples.
Figures show metabolites that significantly changed from day -1 to day 6 within each pathogenic *E. coli* dosing regimen. Points represent individual metabolites, which are color-coded based on the Human Metabolomics Database chemical taxonomy (Wishart et al., 2018). Dotted lines indicate significance cutoffs of $P< 0.1$ (Y-axis) and >2 fold

change in abundance (X-axis); points in the upper left and right quadrants of each graph represent metabolites with significant, >2-fold changes from pre- to post-treatment. Point shape indicates if a metabolite is shared by the acute and chronic treatments (squares) or chronic and control treatments (triangles); circular points indicate metabolites that were either not detected or had non-significant changes in other treatments. Both of the *E. coli*-treated groups exhibited elevated levels of metabolites associated with amino acids and reduced metabolites associated with fatty acids compared to the control group.

Treatment	Metabolite	log2(Fold Change) Compared to Pre-treatment	raw.pval	HMDB Subclass
Control	**glutamine**	**2.9602**	**0.09698**	**Amino Acids, peptides, and analogues**
	N-acetylgalactosamine	2.6842	0.00728	Carbohydrates and carbohydrate conjugates
	succinic acid	2.4595	0.00552	Carbohydrates and carbohydrate conjugates
	alpha-aminoadipic acid	2.2768	0.02794	Amino acids, peptides, and analogues
	hypoxanthine	2.1773	0.00884	Purines and purine derivatives
	methionine sulfoxide	2.0569	0.00225	Amino acids, peptides, and analogues
	lyxitol	1.9588	0.04774	Carbohydrates and carbohydrate conjugates
	arabitol	1.748	0.05888	Carbohydrates and carbohydrate conjugates
	n-acetyl-d-hexosamine	**1.5584**	**0.05627**	**Carbohydrates and carbohydrate conjugates**
	thymine	1.5425	0.07133	Pyrimidines and pyrimidine derivatives
	1,3-diaminopropane	**1.476**	**0.04187**	**Amines**
	xanthine	1.4181	0.00606	Purines and purine derivatives
	mannonic acid NIST	1.4097	0.02947	N/A
	nicotinic acid	1.3675	0.08477	Pyridinecarboxylic acids and derivatives
	benzoic acid	1.3172	0.08262	Benzoic acids and derivatives
	uracil	1.2077	0.00503	Pyrimidines and pyrimidine derivatives
	phosphoethanolamine	1.0487	0.03619	Phosphate esters
	3-phenyllactic acid	1.0408	0.08417	Phenylpropanoic acids
	3-aminoisobutyric acid	**-1.1348**	**0.09078**	**Amino acids, peptides, and analogues**
	oleic acid	-2.0313	0.01224	Fatty acids and conjugates

Color Legend (based on Human Metabolomics Database chemical taxonomy)
Amino Acids, peptides, and analogues
Carbohydrates and carbohydrate conjugates
Amines
Fatty acids and conjugates
Alpha/Beta hydroxy acids and derivatives
Bold= Metabolites in common between chronic and control, or chronic and acute

Table 4. Metabolites Identified by Volcano Plot (Control).
Altered fecal metabolites between pre- and post- treatment samples identified via volcano plot in **Figure 8**. Separated by Treatment and color coded according to provided legend. (Continued on following two pages)

Treatment	Metabolite	log2(Fold Change) Compared to Pre-treatment	raw.pval	HMDB Subclass
Chronic	**putrescine**	**7.3854**	**0.02559**	**Amines**
	galacturonic acid	4.5858	0.06666	Carbohydrates and carbohydrate conjugates
	1,5-anhydroglucitol	4.0501	0.00030	Carbohydrates and carbohydrate conjugates
	3-4-hydroxyphenylpropionic acid	3.9789	0.04041	Phenylpropanoic acids
	N-acetylornithine	3.3798	0.02537	Amino acids, peptides, and analogues
	hydrocinnamic acid	3.173	0.01423	Phenylpropanoic acids
	5-aminovaleric acid	**3.14**	**0.00942**	**Amino acids, peptides, and analogues**
	lactic acid	**3.1051**	**0.00136**	**Alpha/Beta hydroxy acids and derivatives**
	homoserine	3.0972	0.03083	Amino acids, peptides, and analogues
	histidine	3.0914	0.05433	Amino acids, peptides, and analogues
	2-hydroxyglutaric acid	2.9396	0.05149	Short-chain hydroxy acids and derivatives
	3,6-anhydro-D-glucose	2.9151	0.00065	N/A
	n-acetyl-d-hexosamine	**2.885**	**0.04304**	**N/A**
	xylitol	2.878	0.00463	Carbohydrates and carbohydrate conjugates
	indole-3-acetate	2.8719	0.03734	Indolyl carboxylic acids and derivatives
	noradrenaline	2.8607	0.00022	Benzenediols
	3,6-anhydro-D-galactose	2.7422	0.00236	N/A
	elaidic acid	2.7079	0.02820	Fatty acids and conjugates
	ornithine	**2.6345**	**0.00161**	**Amino acids, peptides, and analogues**
	linoleic acid	2.5974	0.03419	Lineolic acids and derivatives
	2,6-diaminopimelic acid	2.5465	0.01058	Amino acids, peptides, and analogues
	glutamine	**2.5433**	**0.00652**	**Amino acids, peptides, and analogues**
	lactamide	2.5408	0.06460	N/A
	3-3-hydroxyphenylpropionic acid	2.4022	0.03186	Phenylpropanoic acids
	isoribose	2.3651	0.00772	Carbohydrates and carbohydrate conjugates
	piperidone	2.3185	0.03358	Piperidinones
	tyramine	**2.3102**	**0.06260**	**Phenethylamines**
	daidzein	2.2622	0.06289	Isoflav-2-enes
	cytosin	2.2317	0.05029	Pyrimidines and pyrimidine derivatives
	conduritol-beta-epoxide	2.1595	0.03229	N/A
	D-erythro-sphingosine	2.1368	0.00623	Amines
	lysine	**2.1064**	**0.00279**	**Amino acids, peptides, and analogues**
	3-hydroxybenzoic acid	2.104	0.01311	Benzoic acids and derivatives
	chenodeoxycholic acid	2.0864	0.01216	Bile acids, alcohols and derivatives
	3-hydroxypropionic acid	1.7908	0.01933	Alpha/Beta hydroxy acids and derivatives
	propane-1,3-diol NIST	1.7861	0.02799	N/A
	glucoheptulose	1.7831	0.03673	Carbohydrates and carbohydrate conjugates
	2,8-dihydroxyquinoline	1.7785	0.02874	Quinolones and derivatives
	2-deoxytetronic acid	1.7763	0.03841	Alpha/Beta hydroxy acids and derivatives
	2-hydroxybutanoic acid	1.6077	0.03709	Alpha/Beta hydroxy acids and derivatives
	N-acetylaspartic acid	1.5787	0.01645	Amino acids, peptides, and analogues
	2,4-diaminobutyric acid	1.5402	0.06848	Amino acids, peptides, and analogues
	1-deoxyerythritol	1.5249	0.05979	Alcohols and polyols
	methyl O-D-galactopyranoside	1.492	0.05124	Carbohydrates and carbohydrate conjugates
	docosahexaenoic acid	1.4534	0.05113	Fatty acids and conjugates
	1,3-diaminopropane	**1.3087**	**0.03344**	**Amines**
	4-hydroxyphenylacetic acid	1.3038	0.06667	1-hydroxy-2-unsubstituted benzenoids
	beta-alanine	1.2516	0.02485	Amino acids, peptides, and analogues
	biphenyl	1.2376	0.09114	Biphenyls and derivatives
	threonine	1.1916	0.02322	Amino acids, peptides, and analogues
	aminomalonate	1.1439	0.03190	Amino acids, peptides, and analogues
	3-aminoisobutyric acid	**1.1261**	**0.07736**	**Amino acids, peptides, and analogues**
	arachidonic acid	1.085	0.03705	Fatty acids and conjugates
	pimelic acid	1.0412	0.01583	Fatty acids and conjugates
	2-aminobutyric acid	-1.1394	0.06930	Amino acids, peptides, and analogues
	pelargonic acid	-1.1737	0.02291	Fatty acids and conjugates
	cis-gondoic acid	-1.3174	0.07020	Fatty acids and conjugates
	lactose	-1.5401	0.03126	Carbohydrates and carbohydrate conjugates
	proline	-2.0382	0.08785	Amino acids, peptides, and analogues
	palmitic acid	-2.1414	0.01565	Fatty acids and conjugates
	beta-glutamic acid	-2.2202	0.02429	Amino acids, peptides, and analogues
	myristic acid	-2.2322	0.03342	Fatty acids and conjugates
	isoheptadecanoic acid NIST	**-2.2658**	**0.03405**	**Fatty Acyls**
	2-ketoisocaproic acid	**-2.3968**	**0.03020**	**Short-chain keto acids and derivatives**
	malic acid	-2.6761	0.06479	Alpha/Beta hydroxy acids and derivatives
	arachidic acid	**-2.8894**	**0.04850**	**Fatty acids and conjugates**
	inositol-4-monophosphate	-3.0858	0.05139	Alcohols and polyols
	stearic acid	**-3.1749**	**0.00923**	**Fatty acids and conjugates**
	alpha-ketoglutarate	-3.2813	0.02381	Gamma-keto acids and derivatives
	fumaric acid	-3.4297	0.09597	Dicarboxylic acids and derivatives
	tocopherol alpha-	-3.7154	0.02734	Quinone and hydroquinone lipids

Table 4 continued. Metabolites Identified by Volcano Plot (Chronic). Altered fecal metabolites between pre- and post- treatment samples identified via volcano plot in Figure **8**. Separated by Treatment and color coded according to provided legend.

Treatment	Metabolite	log2(Fold Change) Compared to Pre-treatment	raw.pval	HMDB Subclass
Acute	lactic acid	4.3107	0.00066	Alpha/Beta hydroxy acids and derivatives
	putrescine	3.9419	0.00796	Amines
	5-aminovaleric acid	3.4838	0.01801	Amino acids, peptides, and analogues
	tyramine	3.4547	0.00134	Phenethylamines
	lysine	2.4848	0.01810	Amino acids, peptides, and analogues
	6-deoxyglucose	2.3381	0.02555	Carbohydrates and carbohydrate conjugates
	ornithine	2.2386	0.05088	Amino acids, peptides, and analogues
	ethanolamine	1.438	0.00705	Amines
	alanine	1.0028	0.06678	Amino acids, peptides, and analogues
	pentadecanoic acid	-1.0604	0.06868	Fatty acids and conjugates
	trans-4-hydroxyproline	-1.6644	0.08314	Amino acids, peptides, and analogues
	heptadecanoic acid	-1.8531	0.00851	Fatty acids and conjugates
	oxoproline	-1.8979	0.09163	N/A
	2-ketoisocaproic acid	-2.3384	0.06232	Short-chain keto acids and derivatives
	stearic acid	-2.9176	0.07804	Fatty acids and conjugates
	arachidic acid	-2.9872	0.00467	Fatty acids and conjugates
	behenic acid	-3.1726	0.06717	Fatty acids and conjugates
	isoheptadecanoic acid NIST	-3.5048	0.00926	Fatty Acyls

Table 4 continued. Metabolites Identified by Volcano Plot (Acute). Altered fecal metabolites between pre- and post- treatment samples identified via volcano plot in Figure **8**. Separated by Treatment and color coded according to provided legend.

Whole genome sequencing

To better understand the gene content of *E. coli* 62-57nal that may contribute to its pathogenicity, its genome was sequenced. The genome of *E. coli* 62-57nal was assembled into 378 contigs of greater than 200 bp each totaling 5.6 Mbp in length and at an average 45-fold coverage. The entire set of 378 contigs was submitted for sequence typing using SerotypeFinder v1.1(Joensen, Tetzschner, Iguchi, Aarestrup, & Scheutz, 2015) and confirmed to be O138 and H14 as reported previously (Ewing, Tatum, & Davis, 1958). Analysis of the assembled contigs by BLASTx (>40% identity, E value <10-5) against a database of known *E. coli* virulence genes identified a number of genes in *E. coli* 62-57nal associated with pathogenesis in both humans and swine, shown in S2 Table. Major identified virulence factors in *E. coli* 62-57nal include hemolysin

(*hlyABCD*), Shiga toxin (*stx2e*), an intimin-like adhesin (*fdeC*), heat-stable enterotoxin II (*stiI*), heat-labile enterotoxin IIA (*eltAB*), an iron scavenging cassette (*chuUAVYTWSX*), and the F18ac+ fimbrial adhesin. Hemolysin (Hly) is an exotoxin that is associated with many pathogenic strains of *E. coli* (Mainil, 2013). Hemolysin is primarily thought to have a role in pathogenesis in extra-intestinal infections, such as those of the urinary tract or septicemia and studies have shown that Hly plays little to no role in clinical signs of diarrhea (Moxley et al., 1998). However, recent data, using both in-vivo murine and in-vitro models, show secretion of Hly can disrupt tight-gap junctions and increase colonic mucosal inflammation (Bucker et al., 2014; Mirsepasi-Lauridsen et al., 2016). This inflammation from Hly has been proposed as a contributing factor for ulcerative colitis in humans (Mirsepasi-Lauridsen et al., 2016). Consistent with *E. coli* 62-57's original edema isolation source, an Stx2e Shiga toxin was identified on a putative prophage element. The Stx2e subtype is known to be associated with edema disease of swine (Moxley, 2000), though in the present study we did not observe signs of edema in challenged piglets. The primary contributors to the observed scouring phenotype were most likely the identified heat-stable enterotoxin II (STIIB) and heat-labile enterotoxin IIA (LT-IIA). While acting by different modes, both STIIB and LT-IIA both have been shown to cause release of water and electrolytes from host membranes, thereby causing diarrhea or scours (Cai, Martínez, Li Qiang Yao, & Yu, 2016 ; W. Zhang et al., 2006). Heat stable and heat labile enterotoxins are the most common exotoxins that are associated with diarrhea in piglets, present in 72% and 57% of ETEC isolates from piglet scours, respectively (W. Zhang, Zhao, Ruesch, Omot, & Francis, 2007). *E. coli* 62-57nal

also contains a number of virulence factors associated with colonization and survival within the host, including the iron scavenging *chu* genes and various adhesin genes coding for the proteins AIDA-I autotransporter, F18ac+ fimbrial adhesin, and the intimin-like FdeC (Mainil, 2013; Mills & Payne, 1995; W. Zhang et al., 2007). Taken together, the presence of these virulence factors explains the observed diarrhea/scouring and colonization phenotype. Furthermore, the enterotoxin mode of action that induces this diarrhea also disrupts the Na^+/K^+ gradients (Cai et al., 2016; Das, Jayaratne, & Barrett, 2018), for which amino acid absorption is dependent upon in the gut (Field, 2003), likely explaining the increased amino acid levels observed in the fecal metabolomes of ETEC-challenged animals.

Accesion	AA Length	Product_name via NCBI Prokaryotic Annotation Pipeline	Blastp in NR database		
			Closest match	Percent ID	Hit Accession
locus_tag="C3727_12735"	445	pseudogene (Putative oxygen independent coproporphyrinogen III (chuW))	heme anaerobic degradation radical SAM methyltransferase ChuW/HutW [Escherichia coli]	100	WP_096987543.1
locus_tag="C3727_14970"	574	pseudogene-frameshifted (type III effector)	type III effector [Escherichia coli]	90	WP_096917050.1
PPA50947.1	632	hypothetical protein (putative type III effector)	hypothetical protein [Escherichia coli]	100	WP_064226471.1
PPA51512.1	656	type III effector	type III effector [Escherichia coli]	100	WP_064225807.1
PPA51885.1	123	Heat-labile enterotoxin IIA, B chain	Heat-labile enterotoxin IIA, B chain [Escherichia coli]	100	WP_104013791.1
PPA51886.1	260	Heat-labile enterotoxin IIA, A chain	Heat-labile enterotoxin IIA, A chain [Escherichia coli]	100	WP_104013792.1
PPA51922.1	1417	intimin-like adhesin FdeC	intimin-like adhesin FdeC [Escherichia coli]	100	WP_104013776.1
PPA52210.1	221	EscR/YscR/HrcR family type III secretion system export apparatus protein	MULTISPECIES: EscR/YscR/HrcR family type III secretion system export apparatus protein [Enterobacteriaceae]	100	WP_000071534.1
PPA52215.1	686	EscV/YscV/HrcV family type III secretion system export apparatus protein	MULTISPECIES: EscV/YscV/HrcV family type III secretion system export apparatus protein [Enterobacteriaceae]	100	WP_000482504.1
PPA52222.1	300	adhesin (F18)	adhesin [Escherichia coli]	100	WP_104013766.1
PPA52374.1	670	type III effector	type III effector [Escherichia coli]	100	WP_064225621.1
PPA52952.1	374	glucosyl transferase	DUF1205 domain-containing protein [Escherichia coli]	100	WP_032310753.1
PPA53090.1	523	pentapeptide repeat-containing protein	pentapeptide repeat-containing protein [Escherichia coli]	100	WP_089591880.1
PPA53687.1	256	heme ABC transporter ATP-binding protein	hemin import ATP-binding protein HmuV [Escherichia coli M718]	98	EGI19221.1
PPA53688.1	318	iron ABC transporter permease	iron ABC transporter permease [Escherichia coli]	99	WP_000910402.1
PPA53689.1	207	anaerobilin reductase	anaerobilin reductase [Escherichia coli]	100	WP_064226003.1
PPA53690.1	164	heme utilization cystosolic carrier protein HutX	heme utilization cystosolic carrier protein HutX [Escherichia coli]	100	WP_000020039.1
PPA53691.1	330	hemin ABC transporter substrate-binding protein	hemin ABC transporter substrate-binding protein [Escherichia coli]	99	WP_001081848.1
PPA53692.1	660	TonB-dependent hemoglobin/transferrin/lactoferrin family receptor hemoglobin/transferrin/lactoferrin family receptor	TonB-dependent hemoglobin/transferrin/lactoferrin family receptor	100	WP_064226004.1
PPA53693.1	342	hemin-degrading factor	hemin-degrading factor [Escherichia coli]	100	WP_064226005.1
PPA53982.1	189	type III effector	type III effector [Escherichia coli]	100	WP_074397995.1
PPA53993.1	253	UDP pyrophosphate phosphatase	UDP pyrophosphate phosphatase [Escherichia coli]	100	WP_064225948.1
PPA54029.1	473	type III effector	hypothetical protein [Escherichia coli]	100	WP_064226704.1
PPA54091.1	170	hemolysin-activating lysine-acyltransferase HlyC	MULTISPECIES: toxin-activating lysine-acyltransferase [Enterobacteriaceae]	100	WP_001372478.1
PPA54092.1	1024	RTX toxin hemolysin A	MULTISPECIES: RTX toxin hemolysin A [Enterobacteriaceae]	100	WP_039022831.1
PPA54093.1	707	alpha-hemolysin translocation ATP-binding protein HlyB	MULTISPECIES: alpha-hemolysin translocation ATP-binding protein HlyB [Enterobacteriaceae]	100	WP_000376539.1
PPA54094.1	478	HlyD family type I secretion periplasmic adaptor subunit	MULTISPECIES: HlyD family type I secretion periplasmic adaptor subunit [Enterobacteriaceae]	100	WP_000860467.1
PPA54678.1	418	leucine-rich repeat domain-containing protein	conserved hypothetical protein [Escherichia coli UMN026]	99	CAR12925.1
PPA55232.1	1287	autotransporter adhesin AIDA-I	MULTISPECIES: autotransporter adhesin AIDA-I [Escherichia]	100	WP_001398178.1
PPA55456.1	87	Shiga toxin subunit B	MULTISPECIES: Shiga toxin Stx2e subunit B [Enterobacteriaceae]	100	WP_000738063.1
PPA55457.1	319	Shiga toxin Stx2 subunit A	MULTISPECIES: Shiga toxin Stx2 subunit A [Enterobacteriaceae]	100	WP_001372428.1
PPA55518.1	71	Heat-stable enterotoxin II	enterotoxin [Escherichia coli]	100	WP_015056453.1

Table 5. *E. coli* 62-57nal Virulence Genes.
Putative virulence genes in the trial strain 62-57nal identified via BLASTp against a custom database of known *E. coli* virulence factors, see materials and methods.

2.3. Conclusion

Animal models are imperative for the advancement of knowledge in human disease as well as the development of therapeutic interventions. Rodent models have often been used as disease surrogates; however, these models are at a disadvantage when it comes to accurately representing human diseases and syndromes. Swine more accurately resemble humans in anatomy, genetics, and physiology, making them a more appropriate animal model for human biomedical research (Jiminez et al., 2015). In this study, weaned piglets were challenged with either a single bolus of ~10^9 CFU of a pathogenic STIIB+, Stx2e+, and LT-IIA+ *E. coli* strain or daily doses of ~10^8 CFU of the same strain. Both chronic and acute treatment groups exhibited significant increases in fecal *E. coli* 62-57nal shedding, intestinal *E. coli* 62-57nal colonization and serum IL-6 levels compared to controls (**Figure 4, 5**). Furthermore, both treatments induced similar levels of dysbiosis as measured by targeted 16S qPCR and untargeted metabolomics (**Figure 7**). These alterations of the microbiome and metabolome consisted of significant increases in Bacteroidetes and *Enterococcus*, amongst others, and increased levels of amino acid metabolites within the feces. These findings imply that high acute doses of inoculum, as are often utilized in studies of *E. coli* gastrointestinal disease in pigs, are not necessarily required to establish a disease state, and lower levels of inoculum may yield a comparable disease state important for the study of chronic inflammation or chronic exposure to pathogens. In this study, fecal

calprotectin measurements and histological examination of intestinal sections from challenged animals did not demonstrate characteristics previously associated with EED in humans. Given the lack of blunted villi and fecal calprotectin, the model may be more representative of a general inflammatory bowel disease rather than EED specifically. The detailed model presented here may be utilized to evaluate novel therapies for human enteropathies or post-weaning diarrhea in piglets. Future studies employing the model described here may benefit from an extended trial period, as well as controlled changes in animal nutrition to achieve blunted villi.

2.4. Materials and Methods

Ethics Declaration

All procedures involving animals and their care were approved and monitored by the Animal Care and Use Committee of the USDA Agricultural Research Service (ARS) Southern Plains Agricultural Research Center (SPARC) and the Texas A&M University Institutional Animal Care and Use Committee (IACUC). All approved methods were carried out in accordance with Animal Care and Use Committee of the USDA SPARC (protocol 2016003) and the Texas A&M University IACUC (2015-0372). Materials and Methods, and Results sections are reported and prepared in accordance with ARRIVE guidelines.

Bacterial strains and culture conditions

Escherichia coli strain NCDC 62-57 (O138:K81(B):H14) was originally isolated from swine showing clinical signs of porcine edema disease (Ewing et al., 1958). This

strain was obtained from the ATCC (ATCC 23545). A spontaneous nalidixic acid-resistant mutant (50.0 μg/ml) of this strain was isolated and used throughout this work, and will be referred to as strain 62-57nal. This mutation was stable through multiple consecutive transfers. The bacterium was routinely cultured in LB broth (Bacto tryptone 10 g/L, Bacto yeast extract 5 g/L, NaCl 10 g/L) or on LB agar plates (LB broth plus 15 g/L Bacto agar) aerobically at 37 °C. Samples obtained from animals challenged with *E. coli* 62-57nal were plated on MacConkey agar (Becton, Dickinson and Co., Sparks, MD) amended with 50 μg/ml nalidixic acid (MacConkey+nal).

Whole genome sequencing of *E. coli* 62-57nal

Total DNA was extracted from an overnight culture of *E. coli* 62-57nal using the Qiagen DNeasy Blood and Tissue Kit following the manufacturer's specifications for bacterial cells (Qiagen, Cat No. 69504). Isolated genomic DNA was sequenced on the Illumina MiSeq platform using Illumina V2 500 cycle reagent chemistry generating paired-end 250 bp reads. Reads were quality controlled using FastQC (bioinformatics.babraham.ac.uk), FastX Toolkit (hannonlab.cshl.edu), and assembled using SPAdes 3.5.0 (Bankevich et al., 2012) at k-mer settings of 21,33,55. Contigs <200 bp or with aberrantly low coverage (<8 fold) were filtered from the assembly to yield 378 contigs totaling 5.6 Mbp with ~45-fold average coverage. The resulting contigs were deposited to Genbank under Bioproject/Accession (PRDF00000000), and underwent automated annotation using the NCBI Prokaryotic Genome Annotation Pipeline (Tatusova et al., 2016). Putative virulence factors were identified based on homology using BLASTx of WGS contigs with a custom database of *E. coli* virulence factors

containing *E. coli*-associated proteins contained in mVirDB (Lihong Chen, Zheng, Liu, Yang, & Jin, 2016) and characterized *E. coli* virulence factors (Mainil, 2013). As a control, the genome of the non-pathogenic lab strain of *E. coli* MG1655 (Accession: GCF_000005845.2) was analyzed against the same database. Hits in common from MG1655 and 62-57nal were excluded based on the presumption that they were part of the non-pathogenic *E. coli* gene repertoire. Protein sequences identified in this initial screening were extracted from the 62-57nal genome and manually investigated using BLASTp and InterProScan to confirm conserved domains were intact and putative gene products were approximately full length (Camacho et al., 2009; Finn et al., 2017). The supplied O- and H- antigen serotype provided by ATCC were also confirmed bioinformatically using the SerotypeFinder v1.1 tool located at the Center for Genomic Epidemiology website (Joensen et al., 2015).

Weaned piglet challenge model of *E. coli* 62-57nal

Animals and facilities. A group of 24 weaned barrows approximately 21 days of age (Landrace × Large White, initial mean BW 6.35 kg) were housed at SPARC (College Station, TX). Pigs were randomly assigned to pens (4.634 m^2) with 2 barrows per pen that had solid concrete flooring and was equipped with a nipple waterer, rubber mat, and feeder. Pigs were provided ad libitum access to water and feed; the diet was a standard phase 1 nursery pig pelleted diet (S3 Table) formulated to meet or exceed the National Research Council (2012) recommended requirements of nutrients (Council, 2012).

E. coli 62-57nal challenge trial. The pigs were held 2 d prior to the start of treatment in order to be pre-screened for endogenous enterotoxigenic *E. coli* (ETEC). Pigs were randomly allotted to one of three treatments: non-challenged control, acute challenge (a single dose of ~10^9 CFU), and chronic challenge (a daily dose of ~10^7-10^8 CFU). After random assignment of piglets to pens and treatments, researchers were unblinded in administration of the different treatments, and samples from each treatment were handled and collected uniformly. Each treatment had a total of 4 pens, 2 pigs per pen, for *n*=8 pigs total. All pigs were housed in the same barn, with pens separated by empty pens to prevent cross-contamination between treatments. *E. coli* 62-57nal was cultured overnight (16-18 h) in Tryptic Soy Broth (TSB; Becton, Dickinson and Co.) at 37 ºC with aeration. The acute treatment received a single dose of 6 ml of overnight *E. coli* culture on d 1 and the chronic treatment received a daily dosage of 6 ml of a 10-fold dilution in phosphate buffered saline (PBS, Corning Cellgro) starting on d 1 until the termination of the trial. Pigs and feeders were weighed on d 0, 1, 3, and 6 for calculation of average daily gain (ADG), average daily feed intake (ADFI), and gain to feed (G:F). The d 1 collection of weight data, feces, and blood was approximately 12 h after the initial *E. coli* dose was administered. All animals were humanely euthanized and necropsied on d 7 for collection of intestinal scrapings and sections which were used for analysis of *E. coli* colonization and determination of villi blunting, respectively.

Pre-screening of animals for endogenous phage and ETEC

Prior to the trial initiation (d -1), pigs were screened for endogenous pathogenic *E. coli*, *E. coli* phages, and enteric bacteria capable of growing in the presence of 50

µg/ml nalidixic acid. Briefly, approximately 1 g of feces from each pen was mixed with 9 ml PBS, vortexed until homogenous, serially diluted (10-fold increments), and plated onto both MacConkey agar plates and MacConkey+nal. A chloroformed aliquot of the first sample dilution was plated onto Tryptic Soy Agar (TSA) plates overlaid with 0.5% top agar (5 g Tryptone, 5 g NaCl, 500 ml dH_2O, 0.5% w/v Agar) inoculated with 100 µl from an overnight culture of *E. coli* 62-57nal. MacConkey+nal plates were screened for breakthrough colonies and TSA plates were screened for plaque formation or zones of clearing to determine phage presence. After overnight incubation, 3 colonies were selected from MacConkey plates from each pen and mixed with 150 µl Tris EDTA (TE) buffer. Each sample was boiled for 10 minutes then centrifuged at 8,000 x *g* for 2 min. These colonies were screened via multiplex PCR for: Universal stress protein A (uspA), heat-labile toxin (LTI), heat-stable enterotoxin I (STI), heat-stable enterotoxin II (STII), Shiga toxin type 1 (Stx1) and Shiga toxin type 2 (Stx2)-encoding genes using previously established and validated primers (Osek, 2001; PETER K. FAGAN & DJORDJEVIC, 1999). Positive bands of appropriate size were confirmed using individual primer sets and resultant PCR products were visualized on a 1.5% agarose gel with gel red (Biotium).

Fecal collection and determination of *E. coli* 62-57nal and fecal calprotectin within feces

A representative fecal sample was collected from each pen on d 1, 3, and 6 to determine fecal *E. coli* populations. The samples were collected in individual, sterile 50 ml conical tubes and transported on wet ice to the laboratory. These fecal samples were

processed and diluted in the same manner as described above, and were also spotted (20 µl) on MacConkey agar containing 50 µg/ml nalidixic acid which were incubated for 10-12 h at 37 °C to avoid colonies merging within the spots. Fecal calprotectin was determined by a commercially available porcine ELISA kit (MyBioSource, San Diego, CA) with a minimum detection limit of 6.25 ng/ml with an intra-assay CV of less than 15%.

Blood sampling and serum analysis

Blood samples were collected from 2 pigs per pen on d 6. Blood was collected from the cranial vena cava via a 20 gauge needle and a 10 ml serum vacutainer tube (BD, with clot activator and gel for serum separation). Tubes were inverted and allowed a minimum of 30 minutes to clot. Samples were centrifuged at 1,600 × *g* for 10 min at 2°C, and the separated serum samples were stored at -80 ˚C until analysis was performed. Serum concentrations of interleukin 6 (IL-6) and interleukin 8 (IL-8) were determined via porcine ELISA kits (R&D Systems, Minneapolis, MN). The minimum detection for IL-6 was 18.8 pg/ml and 62.5 pg/ml for IL-8. Assays were conducted as outlined by the manufacturer.

Intestinal sampling and histology

Pigs were humanely euthanized, necropsied, and had samples collected for intestinal histology and *E. coli* populations from the duodenum, jejunum, ileum, cecum, and colon. Segments of the small intestine (duodenum, jejunum, and ileum) and large (cecum and colon) intestine were tied off with plastic zip ties to prevent cross contamination. Adherent bacterial samples were collected by rinsing the intestinal

mucosal surface with sterile PBS, scraping a 2-3 cm section of the surface with a glass microscope slide, and then placing the sample into sterile 15 ml conical tubes containing 4.25 g of 2 mm glass beads (Walter Stern Inc.) and 8 ml of sterile PBS. Tissue scraping samples were vortexed for 5 min at 3000 rpm on a platform vortex mixer to homogenize samples. Cecum and distal colon contents were also collected and homogenized by thoroughly vortexing 0.5 - 5 g sample in 25 ml of PBS. Sample homogenates were serially diluted (10-fold increments) in PBS and spot-plated (20 ml) to MacConkey agar with 50 µg/ml nalidixic acid. All bacterial population estimates were normalized to the initial sample wet weight.

At the time of the necropsy, histopathologists were blinded for both tissue sample collection and examination of fixed tissue for adherent bacteria as well as evidence of tissue inflammation. For histology, the distal end of each intestinal segment (duodenum, jejunum, ileum, cecum and colon) directly adjacent to the section used for bacterial sampling was collected. Samples were positioned onto a 5.08 cm x 5.08 cm cardboard sections and secured with small clips to prevent tissue curling. Consecutive tissue samples were fixed in Carnoy's solution (60% ethanol, 30% chloroform, 10% glacial acetic acid) at a 20:1 ratio for 30-45 min and 10% neutral buffered formalin (VWR Int., Inc.) at a 10:1 ratio for 24 h, followed by storage in 70% Ethanol until further processing. Tissues were trimmed into longitudinal sections of 5 mm width and embedded into paraffin using standard procedures (Blick et al., 2019). After processing, 4 µm sections were placed on slides and stained with hematoxylin and eosin and evaluated histologically in a treatment blinded fashion. The slides were analyzed by a

board-certified pathologist for rod attachment, presence of inflammation, and morphological changes (i.e., villous blunting, epithelial erosion). A total of 12 pigs, 4 from each treatment, were randomly selected for the evaluation of villus length, and well-oriented and intact villi were measured from the lamina muscularis mucosae layer to villus tip.

Analysis of fecal metabolites

Fecal samples collected from the floors of pens on d -1, 1, 3 and 6 were lyophilized and sent to the West Coast Metabolomics Center at the University of California at Davis for untargeted metabolomic analysis. Untargeted GC-TOF profiling was performed following previously published parameters (Fiehn et al., 2008). The resultant dataset is available on Metabolomics Workbench (Sud et al., 2016), under study number ST001041.

16S qPCR

For the purposes of quantifying select bacterial populations, microbial DNA was extracted from 50 mg of lyophilized feces using the Zymo Quick-DNA Fecal/Soil Microbe Kits following the manufacturer's instructions. Five ng of DNA was used to amplify 16S regions of select bacterial groups using Biorad SsoFast EvaGreen Supermix using reaction conditions described previously (Suchodolski et al., 2012) and family/genus/species specific primers described previously (Suchodolski et al., 2012; Walter et al., 2001). qPCR data is reported as $\log_{10}$ of starting 16S copy number per 5 ng of DNA isolated. Specific primer sets used were for Universal, *Faecalibacterium*,

Streptococcus, *E. coli*, *Fusobacterium*, Firmicutes, Bacteroidetes, *Lactobacillus*, *Ruminoccaceae*, and *Enterococcus*.

Statistical analysis

Growth performance along with cytokine and intestinal bacterial population data were analyzed using the PROC MIXED procedure in SAS 9.3 (SAS Institute, Inc., Cary, NC). The model fixed effect was treatment with pen set as a random effect for growth performance, cytokine and intestinal bacterial population data. Fecal samples were collected on a pen basis; therefore pen was not included as a random effect. Calprotectin levels and fecal colony counts were analyzed as repeated measures using the PROC GLIMMIX procedure. Treatment, day, and treatment × day served as fixed effects. Day of collection also served as the repeated measure with pen as the subject. Metabolite data was normalized to sum, mean-centered and divided by the standard deviation of each variable, and analyzed for significant or trending metabolites between treatments using MetaboAnalyst version 4.5(Xia & Wishart, 2016). Comparison of qPCR LOGSQ values was carried out using JMP Version 13 (SAS Institute, Inc.). qPCR treatment means were compared pairwise on a per time point basis; many of the datasets did not pass the Shapiro-Wilk test for normality and were of small sample size, therefore treatments were compared using the nonparametric Wilcoxon Exact Test. qPCR data was also considered using multivariate methods on a pre/post treatment basis using principal component analysis (PCA). Results were considered significant at $P \leq 0.05$ and marginally significant between $P > 0.05$ and $P \leq 0.10$.

Data Availability

Data pertaining to resultant metabolomic analysis is available at Metabolomics Workbench, under study number ST001041. Genomic sequencing data available from Genbank under Bioproject (PRDF00000000).

3. ISOLATION AND CHARACTERIZATION OF THREE NOVEL PHAGES MANGALISTA, MINZHU, AND MUKOTA INFECTING PORCINE SHIGA-TOXIN PRODUCING *E. COLI*

3.1. Introduction

Pathogenic *Escherichia coli* are common pathogens affecting swine and the primary cause of piglet post-weaning diarrhea (PWD) and edema disease (ED), which are serious problems for the swine industry (Fairbrother & Nadeau, 2019; Fairbrother et al., 2007; Francis, 2002; W. Zhang et al., 2007). PWD, also referred to as "scours" is generally caused by Enterotoxigenic *E. coli* (ETEC), while ED is caused by Shiga-toxin producing *E. coli* (STEC) (Fairbrother et al., 2007). Clinical signs of PWD include several days of diarrhea due to heat labile and/or heat stable enterotoxins produced by the ETEC strains during, as the name implies, the post-weaning period. During this period, piglets are particularly vulnerable and can easily spread ETEC to other pigs resulting in mortality rates of up to 25%. ED also manifests during the post weaning period, and is characterized by sudden death, swelling of the eyelids and forehead as subcutaneous edemas develop due to dissemination of Shiga toxin, most commonly of the Stx2e subtype (Fairbrother & Nadeau, 2019). In humans, ETEC and STEC are common causes of traveler's diarrhea and hemorrhagic colitis, respectively (S.-C. Yang, Lin, Aljuffali, & Fang, 2017). Early life exposure to pathogenic *E. coli*, as well as other routinely encountered enteric pathogens, has been shown to correlate with growth stunting in humans and biomarkers associated with environmental enteric dysfunction in

developing nations (Iqbal et al., 2019b). As reliance on antibiotics both within agriculture and human use is waning due to environmental and health concerns over emergence of antibiotic resistance, new methods for controlling and reducing bacterial populations are required (Lhermie, Grohn, & Raboisson, 2016; Llor & Bjerrum, 2014; Van Boeckel et al., 2017). Bacteriophages, the natural viral predators of bacteria, are a promising alternative to antibiotics for both treatment of ongoing infections and prophylactic protection within swine (Cha et al., 2012; Jamalludeen, Johnson, Shewen, & Gyles, 2009; C. Y. Lee et al., 2017).

In this work, we describe the isolation and characterization of three novel phages, Mukota, Minzhu, and Mangalitsa, infecting the pathogenic *E. coli* strain NCDC 62-57 encoding F18ac+ fimbrial adhesin, heat-labile enterotoxin IIA, heat-stable enterotoxin II, as well as Shiga toxin subtype Stx2e. These phages were isolated in the pursuit of reducing the clinical signs of PWD, induced by the host *E. coli* as described in Chapter 2. A comprehensive analysis on the pathogenesis of this strain has recently been described in a piglet model of gastrointestinal disease, Chapter 2 (Boeckman et al., 2021). Ultimately, the evaluation of an oral phage cocktail's efficacy for the control of PWD clinical signs was unsuccessful due to a number of factors affecting the trial, and thus in-vivo data surrounding these phages is limited. However, here we characterize and evaluate the potential of phage Mukota, a T4-like phage, and phages Minzhu and Mangalitsa, phages belonging to a novel family of myophages encoding a phage RNA polymerase and displaying temperature sensitive phenotypes, for the therapeutic or biocontrol of STEC. Particular attention will be paid to the unique temperature

phenotypes of phage Minzhu and Mangalitsa as they may represent novel control methods of STEC at low temperature.

3.2. Results and Discussion

Isolation of Mukota, Minzhu, and Mangalitsa

In pursuit of isolating phages capable of reducing STEC growth and associated symptoms, a number of phages were isolated from a variety of sources including wastewater and swine fecal samples. To enhance plaque formation and consequent chance of isolating novel phage, phage enrichments were frequently plated at both 37 and 30 ℃ to screen for plaques. Using a panel of ETEC and STEC strains, a total of 38 phages were isolated **(Table 6).** Only 3 of the isolated phages, Mukota, Minzhu, and Mangalitsa, infected the NCDC 62-57 STEC strain of interest **(Table 6, red) (Table 7)**. For the purposes of using NCDC 62-57 in a piglet model of gastrointestinal disease (Boeckman et al., 2021), a spontaneous nalidixic acid resistant mutant was isolated and termed NCDC 62-57nal referred to hereafter as 62-57nal. Since only Mukota, Minzhu, and Mangalitsa were capable of infecting 62-57nal, they received the greatest depth of further characterization. Mukota was initially isolated at 30 ℃, which slightly increased plaque formation as it formed "pinpoint" plaques, after which it was subcultured and routinely propagated at 37 ℃. While Minzhu and Mangalitsa were also initially isolated at 30 ℃, neither could be effectively sub-cultured at 37 ℃ and required propagation at 30 and 22 ℃, respectively. It was at this point initial observations were made of the variable plating defects of Minzhu and Mangalitsa at higher temperatures, further

explored below. Mukota, Minzhu, and Mangalitsa were all determined to share myophage morphology with rigid contractile tails, **Figure 10**.

Phage	Isolation Source	Host Strain - Pathotype	Morphology
Still Creek Influent	Waste Water Influent	H10407 - ETEC	n/a
Uknown Activated Sludge	Activated Sludge	H10407 - ETEC	n/a
AsnuS0	Swine Fecal	WAM2317 - ETEC	Myophage
AsnuB	Swine Fecal	WAM2317 - ETEC	Myophage
BrsoKM	Swine Fecal	WAM2317 - ETEC	n/a
BrsoKS	Swine Fecal	WAM2317 - ETEC	n/a
BrcoKS	Swine Fecal	WAM2317 - ETEC	Myophage
BrcoKM	Swine Fecal	WAM2317 - ETEC	Myophage
Mukota	Swine Fecal	62-57nal - STEC	Myophage
BrcoES	Swine Fecal	WAM2317 - ETEC	Myophage
BrcoEB	Swine Fecal	H10407 - ETEC	n/a
OssaE	Swine Fecal	H10407 - ETEC	n/a
PBEpi	Swine Fecal	WAM2317 - ETEC	Myophage
SoFEK	Cattle Fecal	H10407 - ETEC	Siphophage
P1601	Pooled Sample	P1610 - ETEC	n/a
P1604	Pool Sample	P1604 - ETEC	Siphophage
P1610	Pooled Sample	P1610 - ETEC	n/a
LL2-1	CPT Escherichia Phage	H10407 - ETEC	n/a
ASP1610	Swine Fecal	P1610 - ETEC	n/a
ASP1604	Swine Fecal	P1604 - ETEC	n/a
RosK	Swine Fecal	H10407 - ETEC	n/a
RosE	Swine Fecal	WAM2317 - ETEC	MyoPhage
Minzhu	Swine Fecal	62-57nal - STEC	Myophage
PBSP1610	Swine Fecal	P1610 - ETEC	n/a
PBSP1604	Swine Fecal	P1604 - ETEC	n/a
PBSK	Swine Fecal	H10407 - ETEC	Siphophage
PBSK1	Swine Fecal	H10407 - ETEC	n/a
PBSE	Swine Fecal	WAM2317 - ETEC	n/a
D6P1610	Swine Fecal	P1610 - ETEC	n/a
D6P1604	Swine Fecal	P1604 - ETEC	n/a
VMPP1604	Swine Fecal	P1604 - ETEC	n/a
WyKS	Swine Fecal	H10407 - ETEC	n/a
WyKL	Swine Fecal	H10407 - ETEC	n/a
ChCrE	Swine Fecal	WAM2317 - ETEC	n/a
ChCrK	Swine Fecal	H10407 - ETEC	n/a
Mangalitsa	Swine Fecal	62-57nal - STEC	Myophage
ChCrK1	Swine Fecal	WAM2317 - ETEC	n/a
RosK15	Swine Fecal	WAM2317 - ETEC	n/a

Table 6. STEC/ETEC Phage Isolations.
Phages isolated and/or capable of infecting the STEC/ETEC panel, isolation source, host on which it was isolated, and phage morphology if available reported. Color indicates isolation host: Green=H10407, Blue=WAM2317, Red=62-57nal, Orange=P1610, Yellow=P1604.

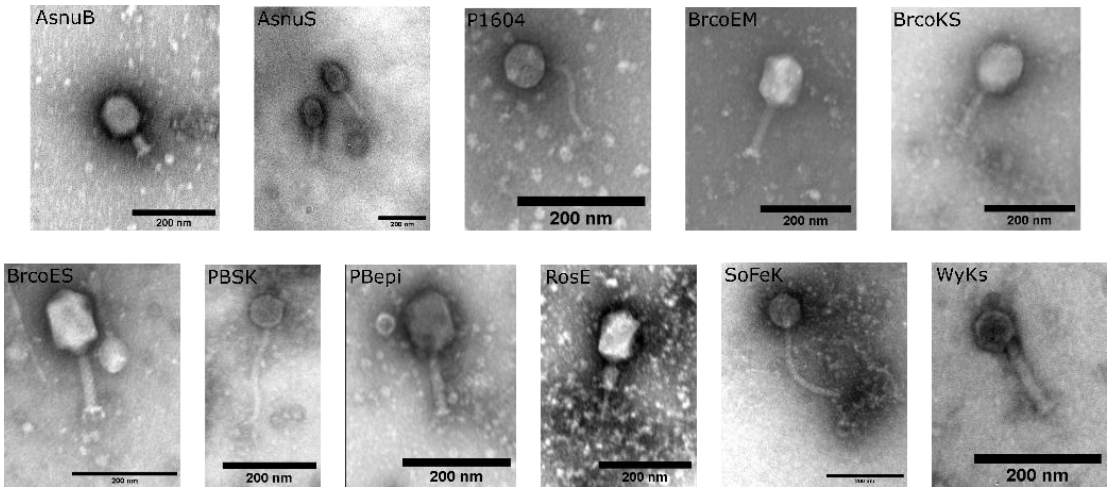


Figure 9. Morphology of Select Phages by TEM.
Phages negative stained with 2% Uranyl Acetate and imaged on JEOL 1200 EX transmission electron microscope at the Texas A&M Microscopy and Imaging Center. Scale bars, 200nm.

	STEC	ETEC			
***E. coli* strain:**	**62-57nal**	**WAM2317**	**H10407**	**P1610**	**P1604**
Source	ATCC 23545	R. Moxley	ATCC 35401	This study	This study
Genome Accession	NZ_PRDF00000000.1	-	NC_000913.3	-	-
Isolation Source	Swine experiencing Edema Disease (EWING WH, 1958)	Piglet with enteric colibacillosis (Moxley, 1998)	Human experiencing acute diarrheal episode (Evans, 1973)	Scouring Piglets	Scouring Piglets
Identified Toxins	α-Hly, STII, LT, Stx2	LT, STII,α-Hly	LT, STI	STI	LT, STII
# of Enrichments	81	84	84	44	44
% of Positives Enrichments	5%	13%	15%	9%	11%
# of Stable Phage Isolates	3	13	12	5	5

Table 7. Summary of Pathogenic E. coli Phage Isolation and Host Strains.

Summary table showing strain designation, associated virulence factors, number of enrichments performed with each strain, % positive enrichments, and number of stable phages isolated.

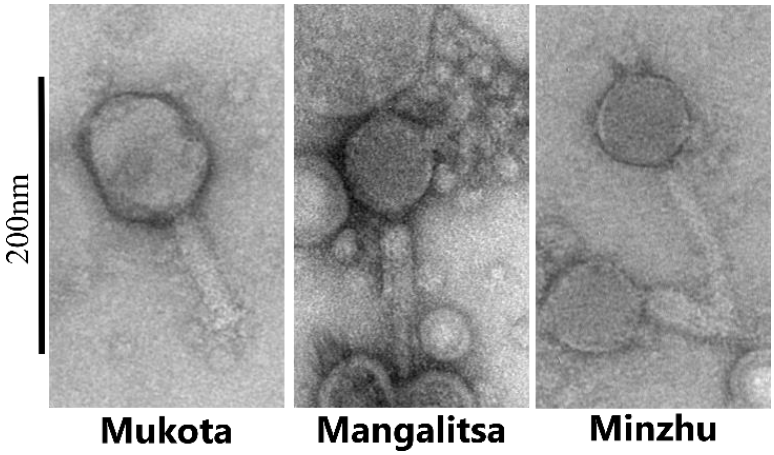


Figure 10. Phages Infecting 62-57nal.
Representative TEM images of the 3 phage isolates infecting STEC 62-57nal. Negative stain with 2% uranyl acetate. Scale bar= 200 nm

Genomes of Mukota, Minzhu, and Mangalitsa

To characterize Mukota, Minzhu, and Mangalitsa, phage genomes were sequenced and annotated. Phage Mukota has a genome of 169,985 bp in length. As can be seen by genomic comparison to phage T4 in **Figure 11,** Mukota is a T4-like phage sharing highly similar genomic content and arrangement. The closest match to Mukota within the NR database of genbank is Escherichia phage HP3 isolated from goose/duck fecal material (Green et al., 2017) with 96% coverage and 97% nucleotide identity. Taxonomically, as per the International Committee on the Taxonomy of Viruses (ICTV), Mukota falls within the genus *Mosigvirus* in the subfamily *Tevenvirinae* within the

family *Myoviridae*. Outside of the T4-like phage proteins, the majority of Mukota proteins absent from the T4 gene repertoire are primarily proteins of unknown function, with the exception of a collagenase-like peptidase and a phospho-2-dehydro-3-deoxyheptonate aldolase. Genomic comparison showed phages Minzhu and Mangalitsa to be related, sharing ~86% nucleotide identity with genomes of 52,329 bp and 53,302 bp, respectively, **Figure 12**. The genomic sequence of Mangalitsa has been reported in brief previously (Atkison et al., 2019). Notably, Minzhu and Mangalitsa are myophages encoding an RNA polymerase, a property frequently restricted to podophages in the family *Autographiviridae* (Lavigne, Seto, Mahadevan, Ackermann, & Kropinski, 2008). In addition to sharing a high degree of nucleotide identity to each other, Minzhu and Mangalitsa had high nucleotide identity with a phiEcoM-GJ1 (Jamalludeen et al., 2008), the founding member of small group of phages in the genus *Carltongylesvirus* characterized by myophage morphology with genomes of ~52kb in size and encoding a T7-like RNA polymerase (P. J. Walker et al., 2020b). Minzhu and Mangalitsa both share ~78% nucleotide identity with phiEcoM-GJ1 and share ~80% with phage ST32, one of the other of few characterized members of *Carltongylesvirus* (Liu et al., 2018) **(Figure 12B)**. As can be seen in **Figure 12**, all four of the *Carltongylesvirus* phages compared share an extremely high degree of genome synteny and nucleotide similarity, with the majority differences being small inserted or deleted ORFs predominantly restricted to a locus of small hypothetical genes directly downstream of the RNA polymerase. The genomic arrangement of the few known members of the *Carltongylesvirus* genus is highly conserved, consisting of the RNA polymerase followed by a conserved stretch of

non-coding regions, and a locus consisting of primarily small hypothetical genes with no known function (**Figure 12, purple shade**). After this region of small ORFs there is a gap in relative coding density, and then a locus of DNA replication machinery including phage encoded DNA primase, DNA polymerase, DNA ligase, and the large terminase (**Figure 12, blue shade**). Finally, the right most arm contains a locus of primarily structural proteins (**Figure 12, green shade**) and a conserved lysis cassette region consisting of holin and endolysin (**Figure 12, pink**); no spanin was identified. Based on genomic analysis, phage Mukota is a fairly typical T4-like phage from the genus *Mosigvirus*, while phage Minzhu and Mangalitsa are more novel phages belonging to the relatively small genus *Carltongylesvirus*, containing RNA polymerases and low-temperature phenotypes (Jamalludeen et al., 2008; Liu et al., 2018).

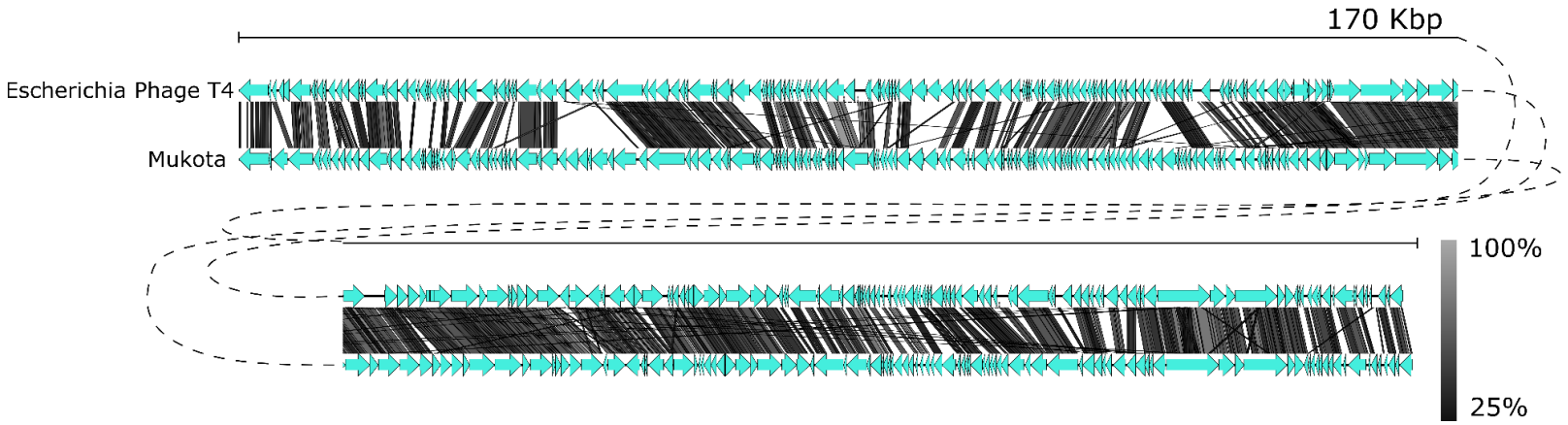
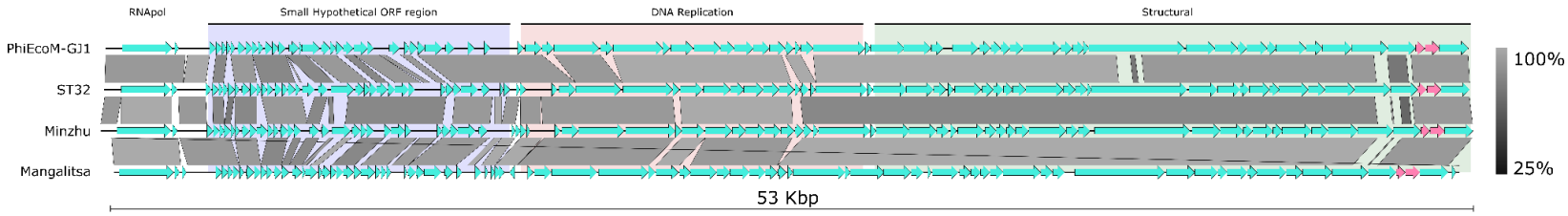


Figure 11. Genome Comparison of Mukota to Escherichia Phage T4.
Mukota is a T4-like phage with conserved genomic content and synteny. Genome map of Mukota (bottom) in comparison to Escherichia Phage T4 (MT984581.1). Genome map generated via EasyFig and comparison lines calculated via BLASTX indicating amino acid similarity. Legend indicates amino acid similarity, genome map split into two levels to assist in visualization.

A



B

	phiEcoM-GJ1	ST32	Minzhu	Mangalitsa
phiEcoM-GJ1	100	84.04	77.92	79.27
ST32	84.04	100	79.11	80.19
Minzhu	77.92	79.11	100	86.4
Mangalitsa	79.27	80.19	86.4	100

Figure 12. Genomic Comparison of Minzhu, Mangalitsa, and Other Members of *Carltongylesvirus*. Mangalitsa and ST32 were compared to two other characterized *Carltongylesvirus* members, the founding member PhiEcoM-GJ1 (NC_010106.1) and ST32 (NC_047830.1). A) Comparison map of conserved genome synteny and nucleotide identity. Shaded regions to indicate 3 major loci with their major functions. Pink, identified lysis genes. Map generated via EasyFig using blastN results. Connecting bars indicate percent nucleotide identity as indicated by legend. B) Overall nucleotide percent identity between indicated genomes. Minzhu and Mangalitsa are more similar to one another than ST32 or PhiEcoM-GJ1.

Receptor of Phage Mukota and Mutations Overcoming Phage Resistance

To further characterize phage-host interactions and evaluate the different sensitivity groups of phages Mukota, Minzhu, and Mangalitsa for their potential use in a phage cocktail a number of spontaneous phage resistant mutants were isolated. Three spontaneous phage-resistant mutants of 62-57nal displaying resistance to phage Mukota were isolated and sequenced, designated 62-57nal-MukotaR2, -MukotaR1, and -MukotaR7. Sequence reads were mapped to the phage-sensitive parental host genome and sequence variants identified **(Table 8)**. The initial phage resistant mutant isolated, 62-57nal-MukotaR2, had an efficiency of plating (EOP) of ~10^{-6} relative to WT 62-57nal. Sequencing revealed a nonsense mutation (Gln30Amber) in the host transcriptional regulator *ompR*, leading to a premature amber stop codon significantly shortening the ORF to 29aa compared to the WT *ompR* product of 239aa (**Figure 13**). OmpR's role in phage resistance was confirmed by introduction of an amber suppression plasmid (pCL15gfpmut2) to 62-57nal-MukotaR2, which restored phage sensitivity and EOP back to ~1 (**Figure 14**). OmpR is a transcriptional regulator (Shimada, Takada, Yamamoto, & Ishihama, 2015), thus unlikely to make physical contact with the phage tail fiber. Instead the phage receptor is likely something under transcriptional control of OmpR. Additional sequencing of other phage resistant mutants -MukotaR1 and -MukotaR7, showed -1 frameshifts (348delC and 19delT, respectively) in *ompC*, an outer membrane protein under transcriptional control by OmpR, and involved in maintaining osmotic homeostasis (Feng, Oropeza, Walthers, & Kenney, 2003). Complementation of -MukotaR1 with *ompC* from the parental host 62-57nal supplied on an arabinose-

inducible plasmid (pBAD24+ompC) restored EOP (**Figure 14**). Furthermore, complementation of the 62-57nal-MukotaR2 containing the *ompR* amber mutation with WT *ompC* also rescued the plaquing phenotype, presumably by bypassing the need for OmpR-mediated regulation of *ompC* and returning the EOP back to approximately 1. Taken together, these results support OmpC as the phage receptor of Mukota.

Additionally, two mutants of phage Mukota capable of overcoming the 62-57nal-MukotaR2 *ompR* mediated resistance showed single nucleotide polymorphisms (SNPs) of long distal tail fiber subunit (His991Arg and His942Arg) were sufficient for overcoming phage resistance (**Table 9**). Given Mukota's similarity to *Escherichia* phage T4, this is perhaps unsurprising given OmpC is also the phage receptor of T4, in addition to LPS (Dawes, 1975; Washizaki, Yonesaki, & Otsuka, 2016). The distal tail fiber subunit of Mukota shares ~60% amino acid identity with the distal tail fiber subunit of T4 (gp37); however, alignment becomes particularly disrupted at the specific residues of gp37 implicated in OmpC binding (residues 932-959 (Islam et al., 2019)) (**Figure 15**). Despite the lack of alignment at these specific residues, the SNPs within the Mukota distal tail fiber shown to overcome host phage resistance are conserved residues within T4 gp37 flanking the region involved in OmpC binding (**Figure 15, stars**). This implies the alteration of these residues is also sufficient to alter the receptor binding domain of the distal tail fiber, at least in Mukota, as is supported by the single amino acid changes being sufficient for the restoration of EOP. Such single missense mutations capable of altering or expanding phage host range have been described previously (Boon, Holtappels, Lood, van Noort, & Lavigne, 2020; Drexler, Dannull, Hindennach,

Mutschler, & Henning, 1991; Le et al., 2013). Given that the EOP of plating is not more severe in the frameshift mutants of *ompC*, -MukotaR1 and -MukotaR7, and that there was still apparent burn through and plaques present, while some represent phage mutants, it is possible phage Mukota also has an inefficient secondary phage receptor. While OmpC as a phage receptor is perhaps common place, mutation of OmpR is a relatively novel mechanism of phage resistance only very recently described elsewhere (Cowley et al., 2018; Mutalik et al., 2020). OmpR is a well characterized transcriptional regulator that through combination of its sensor kinase, EnvZ, regulates expression levels of OmpC and OmpF in response to osmolarity. However recent studies have also shown beyond regulation of *ompC* and *ompF*, OmpR also serves as a global regulator binding many other sites in the bacterial genome (Shimada et al., 2015). Thus, by removal of OmpR, expression of OmpC is diminished to levels approximating the *ompC* frameshift mutants as indicated by similar EOP values shown in **Figure 14 and Table 8**. However, it is likely that repression of OmpC mediated through an *ompR* nonsense mutation likely also suffers additional fitness costs given its role as a global regulator. Particularly, it would be predicted that *ompR* mutants suffer a reduced fitness in response to osmotic stress with limited ability to control OmpC and OmpF expression, in addition to other defects caused by unregulated genes normally controlled by OmpR transcriptional regulation.

Similarly, spontaneous phage resistant mutants of Minzhu and Mangalitsa were also isolated, though sequencing to identify SNPs or indels contributing to phage resistance was inconsistent with a high level of background noise, and thus a reasonable

target for complementation could not be identified. However, phage resistant mutants showed Minzhu and Mangalitsa were in the same sensitivity groups. Tail fibers of Minzhu and Mangalitsa share ~75% amino acid identity and are particularly conserved at the C-terminal ~450aa residues.

Resistant Mutant	Contig.	Mutation	Type	Gene	EOP
62-57nal- MukotaR2	PRDF01000181.1	Gln30Amber	Nonsense	*ompR*	$\sim 10^{-6}$
62-57nal- MukotaR1	PRDF01000220.1	348delC	-1 Frameshift	*ompC*	$\sim 10^{-5}$
62-57nal-MukotaR7	PRDF01000220.1	19delT	-1 Frameshift	*ompC*	$\sim 10^{-4}$

Table 8. Mutations Leading to Mukota Phage Resistance.
Sequencing of phage resistant isolates of 62-57nal. Mutant strain designation, contig of the assembly in which the relevant gene was assembled, the mutation, type of mutation, gene, and EOP of relative to WT 62-57nal.

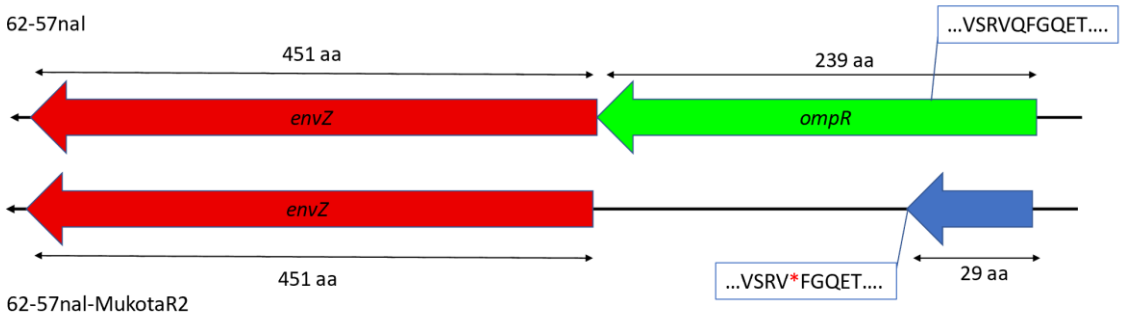


Figure 13. Schematic of OmpR Premature Stop Codon.
Schematic representation of the premature amber stop codon of 62-57nal-MukotaR2. Left in red, *envZ*, ompR's associated sensor kinase. Right, in green and blue, *OmpR*. *OmpR* open reading frame is significantly shortened from 239aa in WT to only 29aa in

62-57nal-MukotaR2. Top WT 62-57nal, bottom 62-57nal-MukotaR2, and amino acid translation of relevant portion.

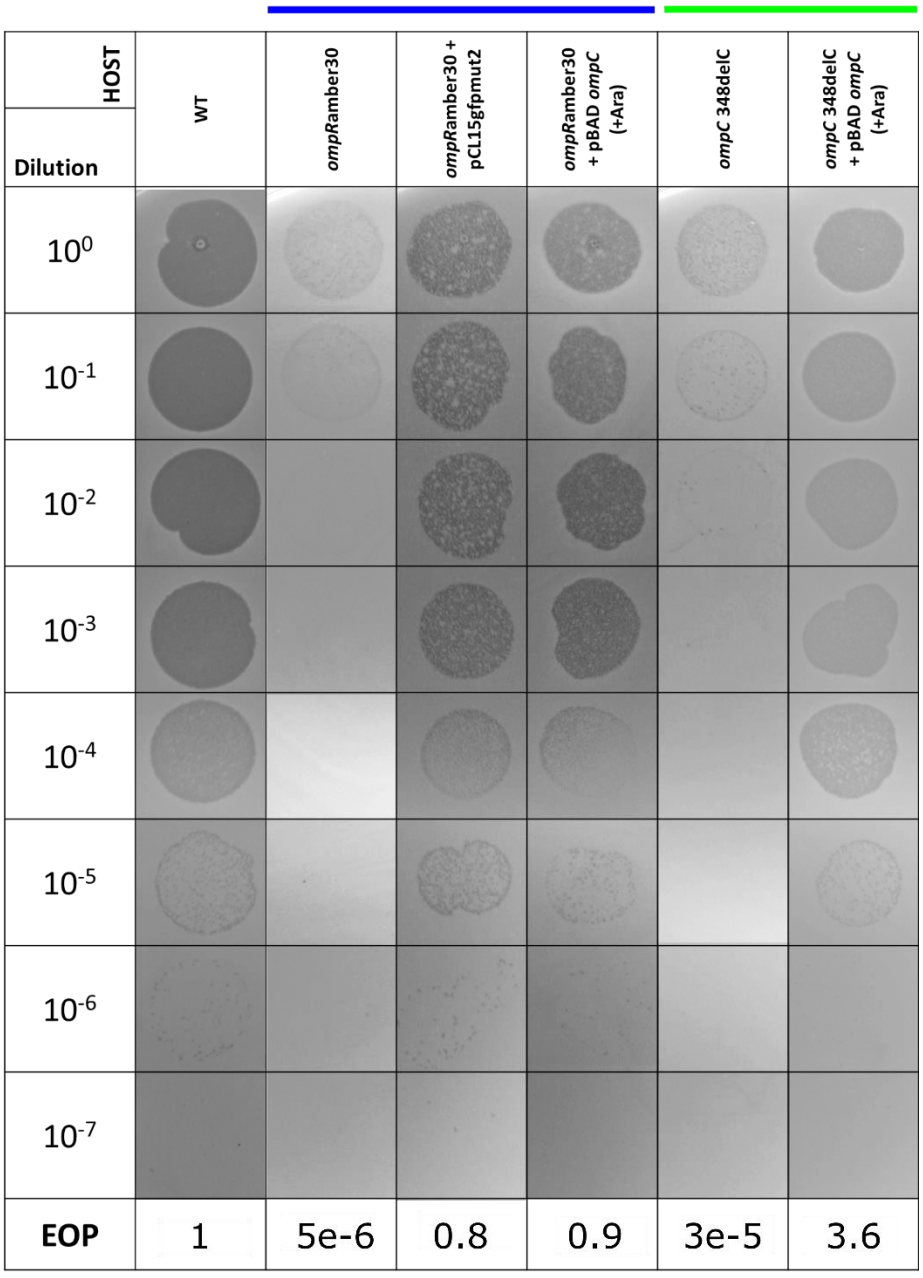


Figure 14. OmpC: the Receptor of Phage Mukota.

Restoration of plating efficiency through Amber stop codon read through via pCL15gfpMut2 or providing WT OmpC via pBAD24 arabinose inducible plasmid. Left, Phage serial dilution. Spot plating of Mukota on WT, resistant isolates, or resistant isolates with respective plasmid as indicated by the column titles. Spot dilutions shown from a representative plating experiment. Color bars above column indicate the same background strain.

Phage	Gene	Mutation	EOP (62-57nal-MukotaR2/WT)
MukotaR2-3	Long tail fiber distal subunit	His991Arg	.7
MukotaR2-4	Long tail fiber distal subunit	His942Arg	.6

Table 9. Mechanisms of Phage Mukota Overcoming Phage Resistance.
Phages capable of overcoming OmpR mediated resistance were isolated and sequenced. Sequencing revealed mutations to the long distal tail fiber subunit. EOP relative to WT 62-57nal and 62-57nal-MukotaR2.

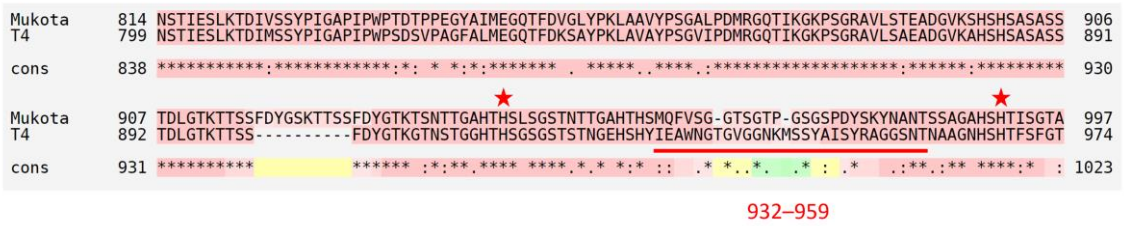


Figure 15. Alignment of Mukota and T4 Long Distal Tail Fiber Subunits.
T-coffee alignment of C-term regions of the distal long tail fiber subunits of Mukota and T4 gp37 (NP_049863.1) to demonstrate conserved amino acid residues. Residues of the T4 reported to be responsible for OmpC binding, 932 to 959 underlined in red (Islam et

al., 2019). Locations of Mukota SNPs overcoming resistance, as reported in **Table 9**, indicated by red stars.

Temperature Sensitivity of Phage Minzhu and Mangalitsa

The inability of phages Minzhu and Mangalitsa to form plaques at higher temperatures was noted at the time of isolation, therefore additional assays were performed to explore this temperature sensitivity. Some phages in the *Carltongylesvirus* genus have been reported to display low adapted temperature phenotypes (Liu et al., 2018). To evaluate temperature effect on plaque formation, serial dilutions of Minzhu and Mangalitsa were plated at different temperatures and EOP's determined relative to the phages preferred plaquing temperatures of 30 and 22 °C, respectively **(Figure 16)**. As can be seen, each phage exhibits plaquing defects at 37 °C, but to varying degrees, and each has a different optimum plaquing temperature. Phage Minzhu has a less severe plaquing defect at 37 °C compared to Mangalitsa, with a preferred plating temperature of 30 °C. Mangalitsa displays no plaques or clearing zones at 37 °C and its optimum plating temperature is at room temperature, ~22 °C. Faint plaques of Mangalitsa can be seen at 30 °C, however these are much more distinct and defined at 22 °C. These results imply Minzhu and Mangalitsa are adapted for optimum performance at temperatures below the classical 37 °C incubation temperature. However, despite their nucleotide similarity, this adaptation to lower temperatures is not uniform between Minzhu and Mangalitsa.

To further characterize this temperature-sensitive phenotype, Minzhu and Mangalitsa were also evaluated in liquid culture. Phage Mukota was also included in

these assays as a point of reference for a more classically lytic phage with growth at 37 °C. At 37 °C, phage Mukota shows potent lytic activity capable of suppressing 62-57nal growth at multiplicity of infection (MOI) values of 1 and 0.1 (**Figure 17, blue**). However, consistent with plating defects observed with Minzhu and Mangalitsa at 37 °C, neither Minzhu or Mangalitsa displayed any ability to suppress growth of 62-57nal at an MOI of 1, and no deflection of growth as compared to the bacteria only control (black) can be seen (**Figure 17, yellow and green, respectively**). At 30 °C, again Mukota displays strong ability to suppress the growth of 62-57nal at all MOI's tested (**Figure 17, blue**). While consistent with its preferred plaquing temperature at 30 °C, Minzhu began to show lytic activity at all MOI's tested showing clear suppression of bacterial growth (yellow) compared to the bacteria only controls (black), though not to the extent of Mukota. Mangalitsa, again consistent with its turbid burn-through of bacterial lawns at 30 °C, still displayed no lytic activity at 30 °C. Finally, at 22 °C, the lowest temperature achievable with this experimental design, the difference between the activity of Minzhu and Mukota at lower temperatures can be can be distinguished. At 23 °C, Mukota is effective at controlling growth of 62-57nal, however only at the highest MOI of 1, while the MOI of 0.1 shows only a mild effect on 62-67nal growth and an MOI of 0.01 shows no difference in bacterial growth. This pattern likely represents control of 62-57nal growth by Mukota at a high MOI through one round of phage replication, after which, as indicated by its performance at lower MOI's, Mukota likely has difficulties effectively propagating at 23 °C. In contrast, Minzhu seemingly performs better at 23 °C at all MOI's tested, even at an MOI of 0.01, indicative of multiple rounds of phage replication

and lysis indicating efficient phage propagation at 23 °C. Mangalitsa, however, still displayed no lytic activity at 23 °C. As with its lower preferred plaquing temperature compared to Minzhu, this may imply that the most efficient replication of Mangalitsa may occur at even lower temperatures. It is also possible Mangalitsa is simply a poor phage at replicating on this strain regardless of temperature in liquid assays, although additional experimentation at below ambient temperatures is necessary to further determine this. Thus, the temperature effects seen in plaquing assays were supported by Minzhu and Mangalitsa growth in liquid culture, and these results support these are true temperature sensitivities and not merely temperature dependent artifacts of plaque formation. Together these results imply Minzhu and Mangalitsa, like other members of the *Carltongylesvirus* (Liu et al., 2018), represent a genus of low temperature phages optimized for temperatures below 37 °C. Thus, these low-temperature phages are presumably adapted for predation of *E. coli* outside the context of the intestine at or below ambient environmental temperatures.

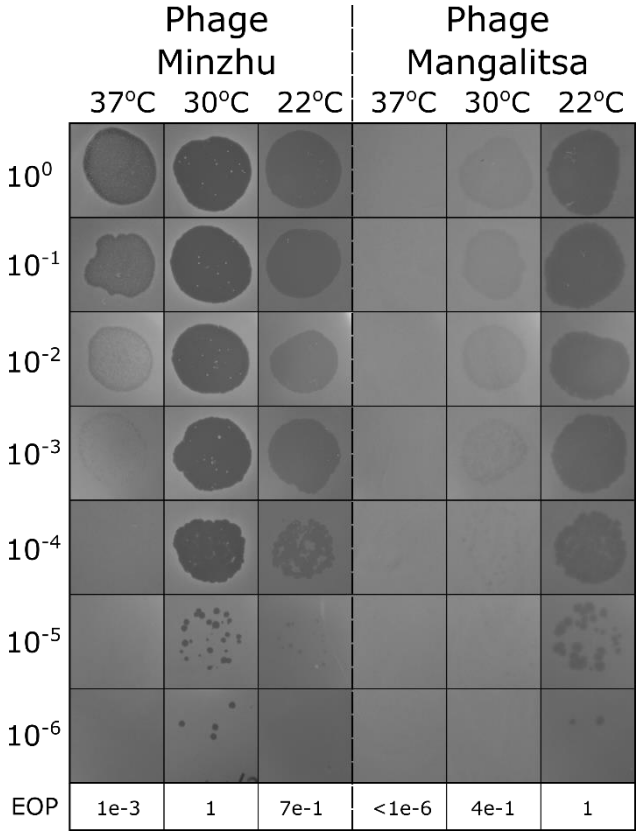


Figure 16. Temperature Dependent Plating Defects of Minzhu and Mangalitsa. EOP of Minzhu and Mangalitsa at indicated temperature. EOP relative to the phage's preferred plaquing temperature, 30 °C and 22 °C for Minzhu and Mangalitsa, respectively.

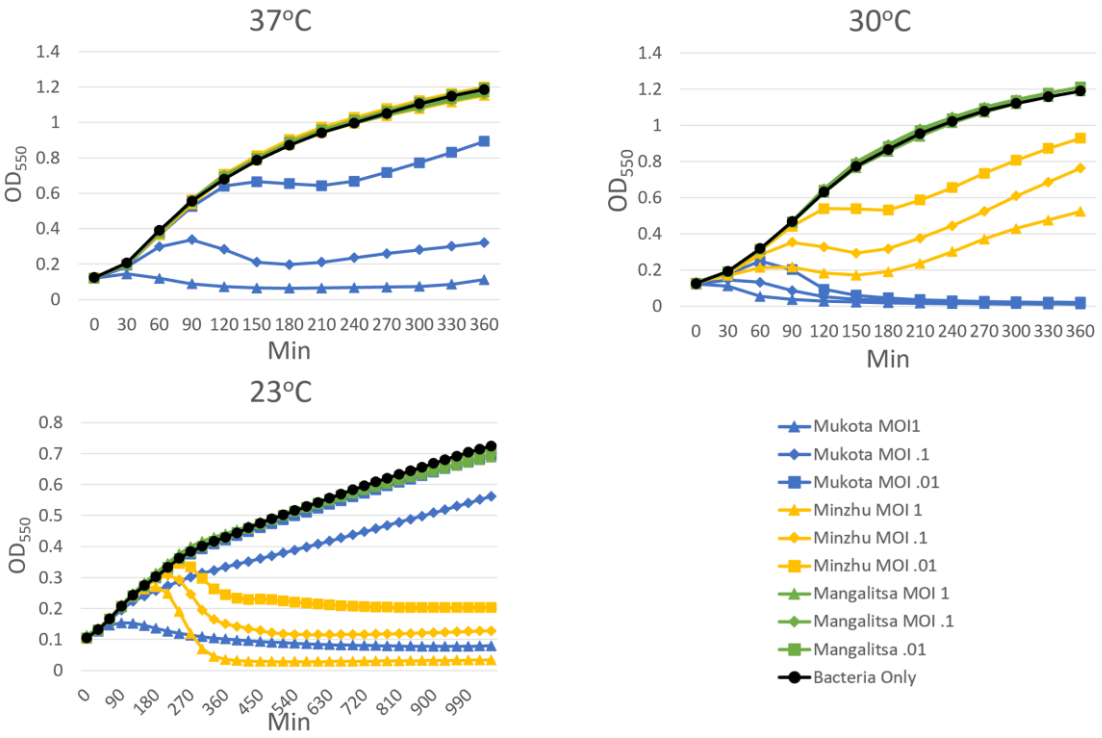


Figure 17. Mukota, Minzhu and Mangalitsa Lytic Activity at 23 °C, 30 °C, 37 °C in Liquid Assays.

Mukota, Minzhu and Mangalitsa were assayed at the three temperatures at multiplicities of infection of 1, 0.1, and 0.01. Phage is indicated by figure legend: Blue-Mukota, Yellow- Minzhu, Green- Mangalitsa. MOI indicated umarker: Triangle – MOI 1, Diamond – MOI 0.1, and Square – MOI 0.01. Mukota shows efficient ability to control 62-57nal at 37 °C and 30 °C, however becomes less effective at 23 °C. Minzhu shows no activity at 37 °C, however lytic activity increases with decreasing temperatures at or below 30 °C. Mangalitsa shows no ability to control 62-57nal growth at any temperature tested.

In-vivo Phage Performance

Ultimately, Mukota and Minzhu were utilized in an attempt to alleviate pathogenic *E. coli* clinical signs caused by an acute dose of NCDC 62-57nal as described previously (Boeckman et al., 2021) (Chapter 2). In brief, an acute dose of NCDC 62-57nal was provided via oral gavage as described in Chapter 2. However, instead of the chronic dosing regimen as described in Chapter 2, a phage cocktail (Mukota and Minzhu) was provided by oral gavage on both day 1 and day 3 as the alternative treatment group. However, for a variety of reasons, including contamination of the control group, the results on efficacy of this phage cocktail were inconclusive. Regardless of the efficacy of phage at alleviating STEC symptoms, it could be seen through tracking of phage levels within swine feces that Mukota roughly followed the levels of the bacterial host that could be recovered from the feces at levels of approximately 10^6 pfu/g of feces three days after the last addition of phage by oral gavage **(Figure 18)**. Interestingly, despite severe plating deficiencies and little to no lytic activity in liquid assays at 37 °C, phage Minzhu was also maintained and recoverable three days post phage treatment, albeit at much lower pfu/g feces compared to Mukota. However, Minzhu was not detectable within the colon contents on day 7, the day of necropsy. The lack of detectable Minzhu from the colon contents may be an indication that Minzhu was inefficient at replication within the piglet intestine given it was not recovered from the colon contents. While the other samples in which Minzhu was detected were fecal samples collected from the pen floor, and thus Minzhu may have been more actively propagating at the ambient temperature of the barn. Given the temperature sensitivity of phage Minzhu, phage Mukota may have benefitted from

alternative phage cocktail partners not showing defects at 37 °C, the internal temperature of the intestine. These results coupled with the plating and liquid assays imply that Minzhu, Mukota, and other low temperature phages may be poor choices for use in therapeutic applications. While the sustained detection of Minzhu from the feces point to the complexities of the in-vivo interactions of phage, host, and the mammalian intestine that standard liquid and soft-agar overlay assays do not fully capture, including phage replication outside of the mammalian host.

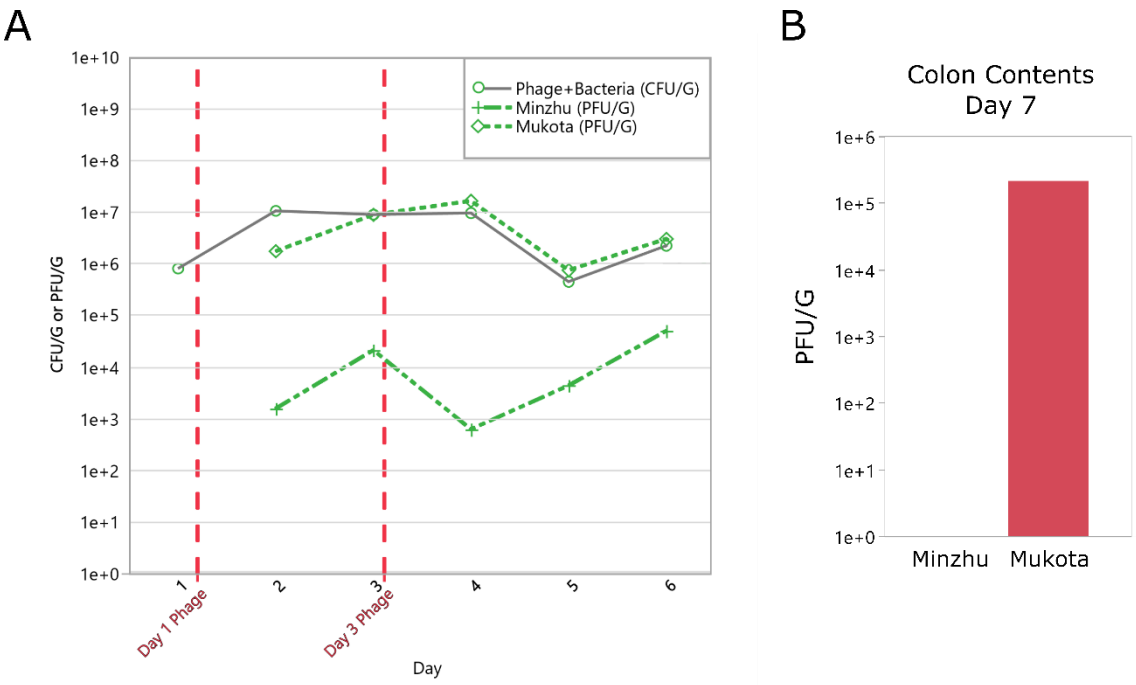


Figure 18. In-Vivo Bacteriophage Survivability.
A) Following an acute dose of 62-57nal, two doses of a phage cocktail consisting of Mukota (at 10^{10} pfu/ml) and KS23M (at 10^{8} pfu/ml) were supplied by oral gavage. Shedding of bacterial strain 62-57nal and phage tracked through direct plating of fecal samples. Solid line indicated recovered bacterial level in CFU/G feces, Grey-Acute dose

with phage. Dashed lines indicate level of recovered bacteriophage. Green diamond dashed line-Mukota. Green plus signs-Minzhu. B) Minzhu and Mukota recovered from colon contents on day 7 (day of necropsy). No Minzhu was detected within the colon contents.

3.3. Conclusion

In this work, we describe the isolation, characterization, and evaluation of three STEC phages infecting an F18ac+ porcine *E. coli* strain NCDC 62-57nal carrying heat-labile enterotoxin IIA, heat-stable enterotoxin II, and Stx2e Shiga toxin for applied purposes. Phage Mukota being a fairly standard T4-like phage utilizing an OmpC phage receptor, displayed strong lytic activity, and ability to control STEC growth in both soft agar and liquid assays. These characteristics make Mukota a good candidate for therapeutic reductions of STEC growth and reduction of *E. coli* induced diarrheal symptoms in humans or in swine production, with appropriate phage cocktails partners. Phages Minzhu and Mangalitsa are novel phages belonging to the relatively small genus *Carltongylesvirus*, containing RNA polymerases and exhibiting low-temperature phenotypes (Jamalludeen et al., 2008; Liu et al., 2018). While Mukota displayed potent lytic activity at 37 °C, Minzhu and Mangalitsa both showed little activity at higher temperatures (**Figure 17**). Based on the EOP and liquid assays, Minzhu appears to be adapted for growth at 30 °C and below, while Mangalitsa seems to be even more sensitive to higher plating temperatures and its optimum replication temperature is presumed to be below 23 °C, although confirmation of its ability to function in liquid culture is required **(Figure 16 and 17)**. The fact that 3 of the 5 characterized

Carltongylesvirus, those examined in this work and ST32 (Liu et al., 2018), displayed a preference for low temperatures suggests this may be a feature of this class rather than an abnormality. ST32 was reported to have significantly greater burst sizes at 20 °C compared to 30 °C or 37 °C (Liu et al., 2018). The initial founding member, phiEcoM-GJ1, was reported to be difficult to propagate to high titer and authors did not attempt lower temperatures (author correspondence), perhaps indicating 37 °C was not the optimum temperature for phiEcoM-GJ1. Most recently, another member of the *Carltongylesvirus, Escherichia* Phage Flopper was reported from a mass isolation of *E. coli* phages in Danish wastewater (Olsen, Forero-Junco, Kot, & Hansen, 2020), though no indication of temperature sensitivity or lack thereof was reported. There also seems to exist a trend for members of *Carltongylesvirus* to be isolated from swine fecal samples and/or infect porcine isolates of ETEC (Jamalludeen et al., 2008; Liu et al., 2018). Assuming these "low-temperature" phages are adapted for predation of *E. coli* in the environment where ambient temperatures are considerably below 37 °C, this trend may be a reflection of the heavily soiled swine housing environment in which these phages may thrive. It stands to reason that there may exist a whole suite of phages adapted to prey on their respective hosts outside of rapid growth environment of the intestine and therefore require adaptations for efficient propagation on slow growing cells or cells in stationary phase. The "low-temperature" phenotype of *Carltongylevirus* coinciding with the unique feature of a myophage with T7-like RNA polymerase, may suggest this RNA polymerase contributes to the low-temperature phenotype, however additional study is required to validate this. Given the genetic similarity of Minzhu and Mangalitsa but with

distinct temperature phenotypes, Minzhu and Mangalitsa may represent an experimental system in which genes contributing to low temperature phenotypes can be readily elucidated.

These temperature defects at 37 °C make Minzhu and Mangalitsa poor candidates for inclusion in phage therapy cocktails. However, as has been suggested of related low-temperature phage (Liu et al., 2018), Minzhu and Mangalitsa may be prime candidates for efficient biocontrol of pathogenic *E. coli* at the lower temperatures often associated with food handling, transport, and storage. Most lytic phages require rapidly growing hosts for efficient replication, and thus existing phage products for biocontrol are hindered by metrics that would assume phage replication on the food products, such as protecting against recontamination (C. D. Carter et al., 2012). If Mukota is exemplary of a standard T4-like lytic phage, this requirement for rapidly growing host cells is demonstrated in **Figure 17** at low MOI's at room temperature. Thus, this represents a gap in existing phage biocontrol formulations and inclusion of such "low-temperature" phages like those identified in this work and other *Carltongylesvirus* may increase the efficiency of phage for control of pathogens at lower temperatures. However, further study is required to elucidate this low temperature phenotype and such explorations of this phenotype may serve to increase the efficacy of biocontrol at or below ambient temperatures through alteration of phage cocktail formulations or the engineering of other lytic phages to perform better at low temperatures.

3.4. Materials and Methods

Bacterial Strain:

The bacterial strain used in this study was Shiga-toxigenic *Escherichia coli* strain NCDC 62-57 (O138:K81(B):H14) originally isolated from swine showing porcine edema disease (Ewing et al., 1958), and obtained from ATCC (ATCC 23545). A spontaneous nalidixic acid resistant mutant was generated and used for selection purposes. The strain was routinely cultured on LB broth (Bacto tryptone 10 g/L, Bacto yeast extract 5 g/L, NaCl 10 g/L) and LB plates (LB broth plus 15 g/L Bacto agar) at 37 °C aerobically. The strain tested positive by PCR for virulence genes encoding Stx2e, Heat Stable enterotoxin 2b and Heat labile Toxin, using methods described in Chapter 2 (Boeckman et al., 2021). Additional virulence factors were identified via BLASTp against a known *E. coli* virulence gene database as described in Chapter 2 (Boeckman et al., 2021). The NCDC 62-57 nalidixic acid mutant was also sequenced using previously reported methods Chapter 2 and is available under accession number (PRDF00000000). NCDC 62-57 phage resistant mutants were isolated by plating with a high multiplicity of infection (MOI) and collecting any surviving bacterial colonies after O/N incubation. Selected mutants were then sub-cultured to ensure purity, efficiency of plating (EOP) evaluated, and sequenced using methods provided in Chapter 2.

Phage Isolation:

Phages were isolated by classical phage enrichment and top-agar overlay techniques (M.H., 1959). In brief, environmental samples in the form of either untreated water, soil, swine feces, or swine barn swabs were collected. Solid samples were added to medium at an approximate ratio of 1:4 (w/v), homogenized, and incubated with LB media for two hours to allow any phage present to diffuse into the media. Samples were

then spun down at 8,000 rpm for 10 min and supernatant collected. Resulting supernatants were then sterilized by the addition of 0.1-1% chloroform by volume. Liquid samples were sterilized by filtration through a 0.22 μm filter. 10 ml of the resulting processed sterilized sample was then added to 40 ml of LB and inoculated with a 100 μl of O/N host culture and incubated O/N with aeration at 37 °C. 10 ml of the resulting enrichment was collected, spun at 8,000 rpm for 10 min, and the resulting supernatant was filter sterilized using 0.22 μm filter. This filtered supernatant was then spotted onto LB plates overlaid with 0.5% top agar (5 g Tryptone, 5 g NaCl, 500 ml dH2O, 0.5% w/v Bacto Agar) inoculated with 100 μl of O/N culture, as well as full plates consisting of 100 μl of O/N culture and 100-500 μl of sterile enrichment media. O/N plates were incubated at 30 and 37 °C then examined for any plaques or burn through.

Phages capable of overcoming 62-57nal-MukotaR2 were isolated via selecting and subculturing phage plaques of Mukota capable of plaque formation on 62-57nal-MukotaR2. After subculture, isolates displaying an equivalent EOP on both WT 62-57nal and 62-57nal-MukotaR2 were considered capable of overcoming OmpR/OmpC mediated resistance.

Phage Propagation

Phage isolates were routinely propagated at their preferred plaquing temperature, room temperature (22 °C), 30°C, or 37°C on LB plates with 5 ml of 0.5% top agar overlay inoculated with 100 μl of 62-57nal.

Phage Genomic DNA and Bacterial DNA Sequencing

Bacteriophage DNA was isolated from sterile high titer phage lysates using the Promega Wizard DNA clean-up kit as described previously (E. J. Summer, 2009a). Bacterial DNA was isolated using the Qiagen DNeasy blood and tissue kit following manufacturer's instructions for bacterial culture. Genomic library prepared with either TruSeq Nano low-throughput kit, Nextera DNA flex kit, or Swift 2S Turbo DNA library kit and sequenced on an Illumina MiSeq platform generating either 150 bp or 250 bp pair-ended reads. Resulting reads were quality control with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed with FASTX-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), as described previously (Atkison et al., 2019). Resulting reads were then assembled using Spades v3.5.0 and closed via Sanger sequencing of PCR products (Bankevich et al., 2012). Resulting phage genomes were then annotated utilizing the Center for Phage Technology's instance of Galaxy with structural and functional annotation pipelines (Atkison et al., 2019; Ramsey et al., 2020). Bacterial DNA sequencing was performed as described in Chapter 2. Genomic comparison maps were generated via EasyFig (M. J. Sullivan, Petty, & Beatson, 2011) with BLASTn and tBLASTX for Mukota and Minzhu/Mangalitsa, respectively, in the CPT instance Galaxy (Ramsey et al., 2020). Overall percent identity tables were generated via progressiveMauve (Darling, Mau, & Perna, 2010) and Convert XMFA to percent identity table tool in CPT Galaxy with DICE method enabled (Ramsey et al., 2020).

Phage Receptor Identification

To identify SNPs and indels potentially contributing to the observed phage resistant phenotype, WT 62-57nal reads and 62-57nal-Mukota resistant reads were separately mapped to the reference WT 62-57nal contigs using Bowtie2 within the CPT Galaxy interface (Cock, Grüning, Paszkiewicz, & Pritchard, 2013; Langmead & Salzberg, 2012a). Default "pair-ended" parameters were used with the exception of the minimum alignment score to be considered `valid` was adjusted to "L,0,-0.1". On the resulting output, SAM tools:MPileup (H. Li et al., 2009) using the default parameters and BCFtools call (H. Li, 2011) with the variants only option was performed to identify potential SNPs and indels. Only variants with a quality score >180 were considered for further analysis. Variants from the WT reads and mutant reads were cross referenced and duplicate variant calls were removed. This resulted in 18 variants for investigation. Variants were located and investigated via ARTEMIS and BLAST (Camacho et al., 2009; Rutherford et al., 2000). Variants that fell within a CDS, or upstream of CDS in putative regulatory regions, and fell within a gene that could logically be connected to phage resistance were considered.

Phage resistant mutant 62-57nal-MukotaR2 was complemented by transformation with plasmid containing an IPTG inducible AMB suppressing system and GFP signal. 62-57nal-MukotaR1, containing *ompC* frameshift, was complemented in trans with WT ompC cloned from the sensitive WT 62-57nal strain on a pBAD24 arabinose inducible plasmid. WT OmpC was cloned into pBAD24 via PCR for amplification and addition of restriction sites, restriction digestion, ligation, and transformed into NEB® 5-alpha Competent E. coli (High Efficiency) (Catalog# C2987H). Resulting constructs were

ultimately transformed into electrocompetent 62-57nal-MukotaR1 and 62-57nal-MukotaR2 cells.

Liquid Assay of Phage Virulence

An O/N culture of 62-57nal was used to inoculate LB broth at 1:100 and grown to exponential phase with aeration to an OD_{600} of 0.5. This subculture was diluted to an OD_{600} of 0.2 in prewarmed LB broth (~10^8 pfu/ml). 180 µl of this culture was added to 96 well micro plates. Phage stocks of Mukota, Minzhu and Mangalitsa were diluted in lambda dilutant to concentrations allowing for multiplicity of infections of 1, 0.1, and 0.01 once 20 µl of phage dilution were added to 96 well plates resulting in a final $1/10^{th}$ dilution. Plates were then incubated at either 30 or 37 °C in a Tecan Spark 10 M plate reader (Tecan Group Ltd., Mannedorf, Switzerland) plate reader with shaking, and OD measurements at 550nm were taken every 30 min for 12 h. For assay at 22°C, plates were incubated without shaking to reduce heat generated by the rotation mechanism and duration of OD measurements was extended to 17.5 h to accommodate for slower bacterial growth. Reported are the average of three biological replicates performed in triplicate (n=3).

Efficiency of Plating Under Different Conditions

Efficiency of Plating (EOP) was determined under different conditions by serially diluting high titer phage stocks and spotting 20 µl of each dilution onto LB plates with a 0.5% top agar overlay inoculated with O/N culture of 62-57nal. EOP assays were performed in duplicate and reported EOP values are the average of 3 biological replicates. Representative image of a EOP assay shown. Various conditions assayed

being temperature and phage resistant isolates of 62-57nal with or without indicated plasmid.

In-Vivo Phage Cocktail

Phage cocktail for oral gavage was prepared containing Mukota (10^{10} pfu/ml) and Minzhu (10^{8} pfu/ml). Phages were concentrated via tangential flow filtration and exchanged into PBS buffer. Piglets were house and maintained as described in Chapter 2, with the exception of the chronic treatment group replaced by a phage treated group receiving a one-time acute dose of 62-57nal and two doses of phage on day 1 and day 3. Each piglet received 5 ml of phage cocktail by oral gavage on days 1 and 3. Prior to phage cocktail treatment, piglets received 1.5% Sodium Bicarbonate by oral gavage to neutralize stomach acid and assist in phage survival to the intestine. Resulting fecal samples were collected, diluted and plated as described in Chapter 2, with the following modifications. For phage plating, fecal dilutions were handled in the same fashion as the bacterial dilutions with the exception of the addition of chloroform to remove endogenous bacteria. Chloroformed phage dilutions were spotted in duplicate on lawns of WT 62-57nal, as well as a Mukota resistant isolate and a Minzhu resistant isolate, to differentiate total phage versus levels of recovered Mukota and Minzhu phage from feces.

4. ISOLATION AND CHARACTERIZATION OF NOVEL PHAGES INFECTING *DESULFOVIBRIO*

4.1. Introduction

Inflammatory bowel disease (IBD) is a major gut health issue that is estimated to affect over 1 million people in the United States (Dahlhamer, Zammitti, Ward, Wheaton, & Croft, 2016; Kappelman et al., 2007). IBD is a chronic inflammation of the intestinal tract, encompassing two related but distinct disorders, Crohn's disease (CD) and ulcerative colitis (UC). CD is characterized by discontinuous chronic inflammation affecting any segment of the gastrointestinal tract leading to chronic pain and diarrhea (Roda et al., 2020). UC is characterized by inflammation and ulcers often restricted to mucosal layer of the colon, also leading to chronic diarrhea and is often associated bloody stool (Kobayashi et al., 2020). While the exact causes of UC and CD remain elusive, prevailing theories suggest a disruption of gut microbiota leads to dysbiosis and inflammation of the intestine (Guinane & Cotter, 2013; Kobayashi et al., 2020; Roda et al., 2020).

In line with this dysbiosis theory, increased levels of sulfate-reducing bacteria (SRB) have long been associated with IBD, and while it is likely not the sole causal group of microbes, it has been suggested as an initiator of the UC disorder and responsible for maintaining the UC condition (Gibson, Cummings, & Macfarlane, 1991b; Zinkevich & Beech, 2000a). SRB can be carried in colon mucosa of healthy individuals, though studies have reported higher incidence/colonization of SRB in IBD,

particularly UC, patients and/or increased levels of hydrogen sulfide (H_2S) in IBD patients (Coutinho et al., 2017b; Fite, 2004a; Gibson et al., 1991b; Levine, Ellis, Furne, Springfield, & Levitt, 1998b; Loubinoux, Bronowicki, Pereira, Mougenel, & Le Faou, 2002a; Pitcher, Beatty, & Cummings, 2000b; Zinkevich & Beech, 2000a). In addition to their clinical relevance, SRB have also been associated with considerable industrial economic losses as they are also capable of corroding metal and stonework, such as pipelines, and are capable of souring oil fields (Barton & Fauque, 2009; Postgate, 1984). Furthermore, SRB consortia from UC patients cause epithelial cell apoptosis in culture (Coutinho et al., 2017b). H_2S represents a plausible mechanism by which SRB contributes to IBD. H_2S is cytotoxic at high levels and inhibits butyrate oxidation (Singh & Lin, 2015a). Furthermore, IBD has been reported to be more prevalent in westernized countries, where high levels of sulfate are consumed in the form of various food preservatives or high protein diets, in which only a fraction of the sulfate is capable of being absorbed leaving the remainder to potentially fuel H_2S production in the intestine (Hanauer, 2006; Pitcher & Cummings, 1996). Furthermore, IBD relapses and increased levels of SRB are linked with high fat/meat/sulfur diets (Rey et al., 2013; Singh & Lin, 2015a). However, given the slow growth and difficult cultivation associated with SRB, their role within the gut has received limited study and a truly causal or mechanistic link to IBD has yet to be established.

Of the SRB, *Desulfovibrio* spp. are the most predominant in the human gut, with up to 92% of SRB isolated from UC patients belonging to this genus (Gibson et al., 1991b; Pitcher & Cummings, 1996). *Desulfovibrio* is a Gram-negative obligate anaerobe

which offputs H_2S as part of its respiratory process (Postgate & Campbell, 1966) and is innately resistant to many antimicrobials (Y. R. Chen et al., 2019; Dzierzewicz, Cwalina, Jaworska-Kik, Weglarz, & Wilczok, 2001; Lozniewski, Labia, Haristoy, & Mory, 2001; Nakao et al., 2009). Of the *Desulfovibrio* spp., *Desulfovibrio piger* appears to be the most widely carried in humans, with higher instances of *D. piger* associated with active episodes of UC (Loubinoux et al., 2002a; Pitcher et al., 2000b; Zinkevich & Beech, 2000a). In addition to *D. piger*, other clinically-associated species include *D. desulfuricans* and *D. fairfieldensis*, which have also been associated with UC (Gibson et al., 1991b; Loubinoux et al., 2002a). In addition to these connections to IBD, *Desulfovibrio* spp. have also been implicated in rare acute infections and bacteremia (Goldstein, Citron, Peraino, & Cross, 2003; Hagiya et al., 2018). Given these associations with IBD, it is possible that a reduction of *Desulfovibrio* spp. in the GI tract may reduce H_2S production, alleviate IBD symptoms, and help restore gut health.

Therapeutic modulation of the gut microbiome has been suggested as an intervention for chronic inflammatory diseases of the gut (Guinane & Cotter, 2013; McCarville, Caminero, & Verdu, 2016). A suitable form of therapy is required for this modulation: classical antibiotics are ill suited to this role as they are likely to increase levels of dysbiosis (Guinane & Cotter, 2013) and SRB are resistant to many broad-spectrum antibiotics (Nakao et al., 2009). Bacteriophages (phages) are the viruses of bacteria and are among the major predators of bacteria in nature. Phages are highly specific to their hosts, infecting bacterial strains with minimal impact to commensal bacteria (Sheth, Cabral, Chen, & Wang, 2016). In recent years, the use of phages as

antibacterials has received renewed interest and has effectively been used to treat acute bacterial infections (Lin, Koskella, & Lin, 2017; Schooley et al., 2017a; Young & Gill, 2015b). Phages make up an important part of the healthy human gut microbiota and play a role in modulating bacterial populations (Lim et al., 2015; Rodriguez-Valera et al., 2009). The ability of phages to modulate the microbiome has also been demonstrated experimentally in mouse models (Duerkop, Clements, Rollins, Rodrigues, & Hooper, 2012; Reyes, Wu, McNulty, Rohwer, & Gordon, 2013). These features make phages an attractive choice for a targeted reduction of *Desulfovibrio* in the gut.

Prior work exploring *Desulfovibrio* phages has been fairly limited. Since the 1970's less than 15 publications on the topic have been published, and these studies often only examine induction of *Desulfovibrio* prophages. Most recently, *Desulfovibrio* phage work has focused on in-silico examinations of *Desulfovibrio* prophage elements (Crispim et al., 2018). Only a few instances of successful *Desulfovibrio* phage isolation from the environment have been reported. These rare reports of isolation of *Desulfovibrio* phages often report difficulty in plaque formation either due to technical limitations or lack of appropriate plating host, (Crispim et al., 2019; Crispim et al., 2018; Eydal, Jagevall, Hermansson, & Pedersen, 2009; Handley, Adams, & Akagi, 1973; Kamimura & Araki, 1989; Rapp & Wall, 1987; Seyedirashti, Wood, & Akagi, 1991a, 1992; Summer et al., 2011; Summer, Summer, Gill, & Young, 2009; Walker et al., 2006). Furthermore, no isolated *Desulfovibrio* phage genomes exist within Genbank at the time of this writing. In order to explore the diversity of *Desulfovibrio* phages and evaluate their practical potential for phage remediation of IBD, or their use in industrial

settings to inhibit SRB growth, we sought to isolate a panel of *Desulfovibrio* spp. and associated phages to explore their genomic diversity and lytic potential. In this work, we present perhaps the most comprehensive characterization of *Desulfovibrio* phages to date, and report the isolation and characterization of six *Desulfovibrio* phages including genomic analysis, morphology, and evidence of lytic activity. Additionally, given the reported technical limitations, additional focus will also be paid to the method of phage isolation to assist future endeavors in this field.

4.2. Results and Discussion

Desulfovibrio Isolation

To obtain a panel of *Desulfovibrio* strains which bacteriophages could be isolated against, model strains were procured from culture collections and *Desulfovibrio* spp. were directly isolated from a variety of sources. Model *Desulfovibrio* isolates were obtained from cultures collections including *D. vulgaris* (strain Hildenborough) (ATCC 29579), *D. fairfieldensis* (ATCC 70045), and *D. piger* (ATCC 29098). Additional strains were then isolated from the environment following classical SRB isolation techniques (Postgate, 1984) utilizing either Postgate Medium B (PMB) or Modified Barr's medium for sulfate reducers (MB), supplemented with iron for the indictive formation of black iron sulfide, FeS. **Figure 19A** shows representative Hungate tubes either positive (+) or negative (-) for SRB growth. Positive tubes were serially diluted and plated onto indicator plates and individual colonies isolated (**Figure 19B**). **Figure 19C** shows early formation of FeS particles in a positive Hungate tube prior to changing completely black. Screened samples included anoxic mud, activated sludge, wastewater influent,

canine feces, human feces, fresh water creeks, rumen fluid, and rumen contents. Isolates were classified to the genus or species level by sequencing of the complete 16S rRNA gene. 16S sequencing revealed selection of black colonies as has been described in the literature was not sufficient to limit background growth of other bacterial species. A number of other genera showed the capability to generate black colonies on the selective media, such as *Megasphaera*, *Veillonella, Acidaminococcus, Clostridium, Citrobacter, Klebsiella, Escherichia, and Streptococcus* **(Table 10)**. Classical methods with MB or PMB only showed ~7% of colonies isolated were actually *Desulfovibrio* by 16S sequencing. Given this selection rate, additional methods were attempted to differentiate *Desulfovibrio* from non-*Desulfovibrio* species prior to 16S sequencing such as confirmation of the dissimilatory sulfate reductase, known to fluoresce red under alkaline conditions when exposed to long wave UV (Warren, Citron, Merriam, & Goldstein, 2005). **Figure 20** displays such florescence, although this too showed to be unreliable in distinguishing *Desulfovibiro* from other genera as only those indicated by red boxes were the *Desulfovibrio* controls. Media used for isolation was ultimately switched to the EDCM media recently described by Chen et al. (2019). This EDCM medium takes advantage of the innate antibiotic resistances of *Desulfovibrio* by inclusion of vancomycin and colistin sulphate in the medium, drastically reducing background growth of non-*Desulfovibrio* species. A side-by-side comparison of PMB and the new EDCM for isolation of *Desulfovibrio* from the same environmental sample showed EDCM displayed a much higher success rate of *Desulfovibrio* isolation **(Table 11)**. Using PMB, only (1/21) black colonies isolated was a *Desulfovibrio* species,

however with EDCM (14/15) black colonies isolated were a *Desulfovibrio* species. In total, 11 *Desulfovibrio* spp. isolates were collected **(Table 12)**. Representative cell morphology of select isolates can be seen in **Figure 21**. Note each strain displays the characteristic curved rod morphology associated with *Desulfovibrio* spp. Additionally, *D. intestinalis* was also observed in a filamentous form as well as curved rods, such filamentous growth has been described previously (Postgate, 1984).

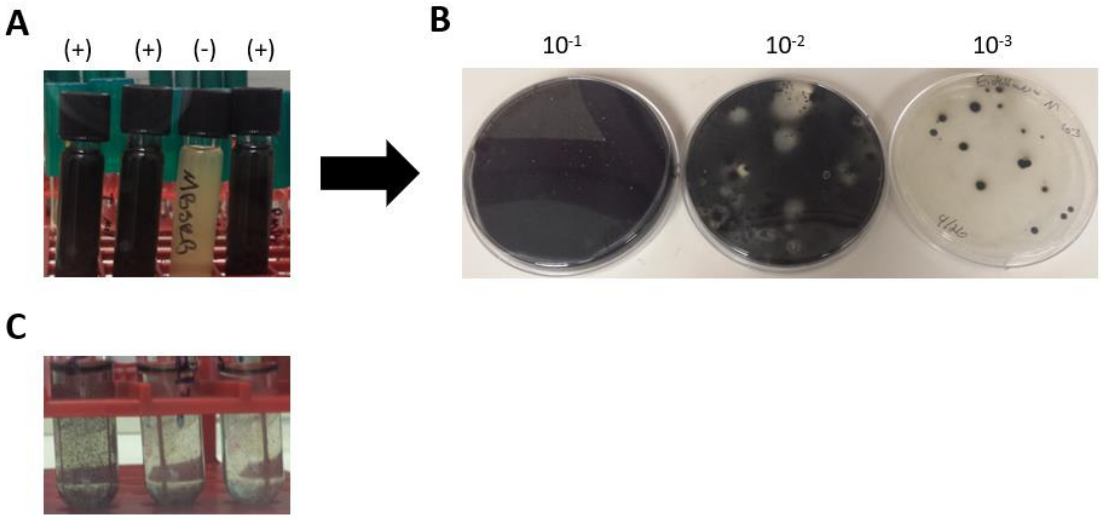


Figure 19. Isolation of *Desulfovibrio* Isolates.
A.) Representative image displaying hungate tubes positive (+) for SRB growth as indicated by formation of FeS and negative (-) tubes. B.) Positive tubes serially diluted 10-fold and plated on plates supplemented with Fe to obtain individual black colonies. C.) Example of early FeS particle formation in liquid culture, prior to entire tube becoming completely black.

Desulfovibrio Isolation success rate by classical methods	
	MB/PMB
(*Desulfovibrio*)/(Number of Black colonies isolated)	2/27
Success rate	7.41%
Non-*Desulfovibrio* sp. also isolated	*Megasphaera* *Veillonella* *Acidaminococcus* *Clostridium* *Citrobacter* *Klebsiella* *Escherichia* *Streptococcus*

Table 10. Success Rate of *Desulfovibrio* Isolation by Classical Methods.
Number of *Desulfovibrio* isolates obtained by classical methods with either MB or PMB media. Also indicated are background species obtained also capable of forming black colonies.

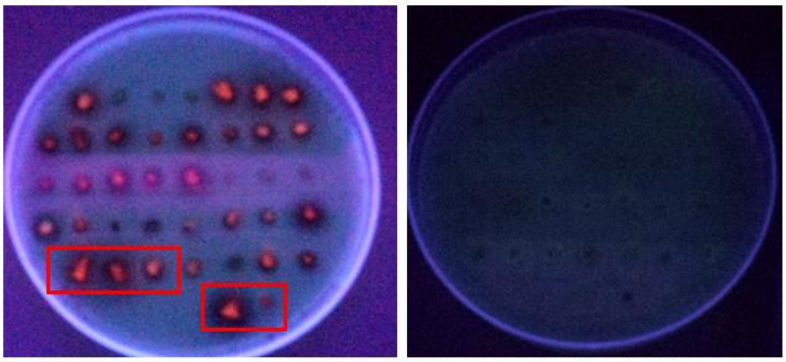


Figure 20. Selection Attempt of *Desulfovibrio* Species by Desulfoviridin.
Attempt to further select for *Desulfov*ibrio species through fluorescence of dissimilatory sulfite reductase gene also known as desulfoviridin. Both panels, exposure to long wave UV (365 nm). Left panel, red/orange fluorescence obtain by flooding a patch plate of

putative *Desulfovibrio* isolates with 2M NaOH. Red boxes indicate control *Desulfovibrio* species. Right panel, no fluorescence observed without presence of NaOH.

	PMB	EDCM
Total	1/21	14/15
Percent	4.76%	93.33%
Non-*Desulfovibrio sp.*	*Escherichia* *Pseudomonas* *Shewanela algae* Unidentified bacterium *Enterobacter*	*Escherichia*

Table 11. Direct Comparison of PMB and EDCM Media.
Side-by-side comparison of PMB vs EDCM for selection of *Desulfovibrio* from the same environmental sample. Also indicated are, non-*Desulfovibrio* species isolated from the respective selection media PMB and EDCM.

Desulfovibrio Panel

Species	Source	Strain	Environmental or Mammalian Isolate
D. vulgaris	ATCC 29579	Hildenborough	Environmental
D. piger	ATCC 29098	VPI C3-23	Mammalian (Human)
D. fairfieldensis	ATCC 70045	FH26001/95	Mammalian (Human)
D. vulgaris	Anoxic Marsh Mud	Galveston #10	Environmental
D. vulgaris	Anoxic Marsh Mud	Galveston #12	Environmental
D. vulgaris	Anoxic River Mud	FM521 #22	Environmental
D. vulgaris	Anoxic River Mud	FM521 #23	Environmental
D. vulgaris	Anoxic River Mud	FM521 #30	Environmental
D. desulfuricans	Anoxic Creek Mud	Edelweiss	Environmental
D. desulfuricans	Anoxic River Mud	FM521	Environmental
D. desulfuricans	Anoxic River Mud	FM521	Environmental
D. intestinalis	Fecal Sample	JB1	Mammalian (Human)
D. legalli	Fecal Sample	CL#1	Mammalian (Cattle)
D. legalli	Fecal Sample	CL#11	Mammalian (Cattle)
Unidentified SRB	Fecal Sample	RS#1	Mammalian (Swine)

Table 12. Summary of *Desulfovibrio* Panel.
Summary of the *Desulfovibrio* isolates that comprised our *Desulfovibrio* collection, including species as determined by 16S PCR, isolation source, and strain designation.

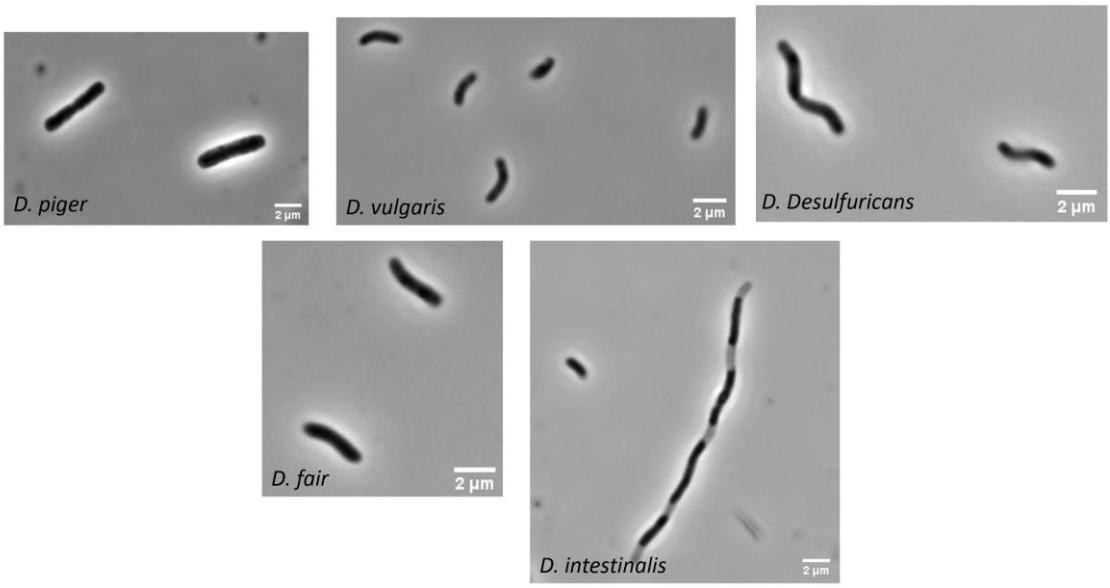


Figure 21. Representative Morphology of Select *Desulfovibrio* Isolates.
Phase contrast microscopy to determine characteristic morphology of *Desulfovibrio* isolates, which can be curved or straight rods, spirilloid, or display filamentous growth (Postgate, 1984).

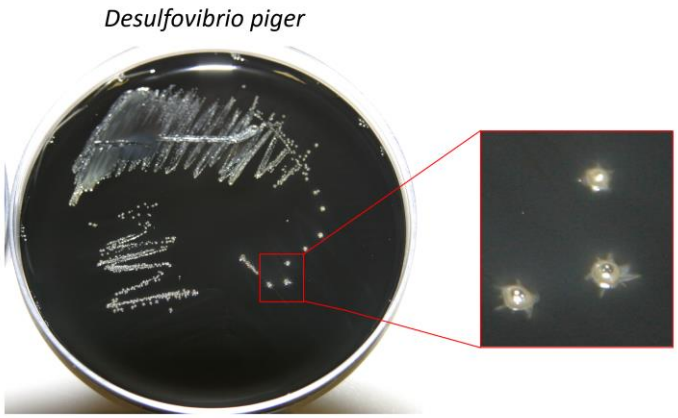


Figure 22. Representative Colony Formation of Individual *Desulfovibrio* Colonies.
Representative colony formation of *D. piger* on PMB agar plates supplemented with Fe^{2+}. Individual colonies initially present as black but with time take on a silver sheen.

Desulfovibrio Phage Isolation

Using the panel of *Desulfovibrio* isolates described above, classical phage isolation methods (Adams, 1959) were employed with the following modifications in efforts to isolate novel phages capable of infecting *Desulfovibrio*. Given the reports in the literature of difficulty or technical issues regarding phage plaque formation, additional focus will be given to methods of *Desulfovibrio* lawn formation. Plaque formation via soft agar overlays is foundational to bacteriophage work; therefore a number of modifications were made to the media to be conducive to phage work, described in the materials and methods. Presence of an opaque precipitate is a common feature of media for SRB growth. This precipitate can inhibit assays which seek to measure optical density, though this is circumvented by utilizing variations of the classical media most of which contain sodium citrate, in order to chelate cations preventing precipitate formation. This is an obstacle to *Desulfovibrio* phage work as citrate is a known inhibitor of phage infection (Shafia & Thompson, 1964), and thus citrate was omitted from media intended for use with phage. The omission of citrate, along with additional modifications detailed in the methods, resulted in the modified medium used for phage isolation and propagation, designated Modified Barr's for phage, or MBfP. MBfP yielded a clear medium without citrate, optionally supplemented with iron(II) sulfate heptahydrate (3.6 mM final concentration). The methods described here yielded a confluent *Desulfovibrio* soft agar lawn conducive to plaque formation and

visualization. In brief, for the isolation of phage sterile environmental samples were combined with pre-reduced MBfP medium and allowed to become anoxic in the anaerobic chamber for ~2 h before addition of $1/10^{th}$ volume rapidly growing *Desulfovibrio* culture grown to an approximate OD_{600} of 0.3 - 0.5 (24-48 h old). Enrichment culture was then incubated for 48-72 h before being centrifuged and supernatant collected and filter sterilized through a 0.22 μm PES filter. Resulting supernatant was then spotted onto soft agar lawns, inoculated with a concentrated cell pellet of respective *Desulfovibrio* host on MBfP agar plates supplemented with Fe^{2+}. Soft agar overlays were poured on a steel plate prechilled to -20 °C to facilitate solidification of top agar lawns in the 37 °C anaerobic chamber environment. Enrichments were also screened by liquid assay and monitored for inhibition of growth via OD as an alternative method in case plaques were unreliable. Liquid cultures displaying inhibited growth were filter sterilized. Filter sterilized supernatant was then further screened for plaque formation and/or diluted and screened again for inhibition of growth in liquid culture until plaques could be obtained. Top-agar overlays were incubated for 16-24 h, at which point plates appear black, and were removed from the anaerobic chamber and exposed to the atmosphere for a period of 6-8 h until FeS begins to dissipate (**Figure 23**). It was found addition of Fe^{2+} to the agar plate greatly aided in visualization of plaques, as can be seen (**Figure 23 and 24**). In the presence of Fe^{2+} following the method described here, phage plaques retain their FeS black color for a longer period of time relative to the surrounding uninfected lawn allowing enhanced visualization. Without iron supplementation, as can be seen in **Figure 23 and 24**, serial

phage dilutions often only result in opaque burn-through making plaque isolation and visualization difficult.

Using these methods, six *Desulfovibrio* phages were isolated from wastewater and freshwater stream samples (**Table 13**). Additionally, a notable T7-like functional prophage element, ProddE, was also identified in *D. desulfuricans* strain Edelweiss further described in Chapter 5. Notably, all phages isolated only infected a single strain, *D. desulfuricans* Edelweiss. This may represent simple differences of surface receptors on different *Desulfovibrio* strains or differences in respective prophage content and/or anti-phage defense systems. It should be noted however, that our panel of *Desulfovibrio* isolates only contained two other *D. desulfuricans* strains, and while *Desulfovibrio* spp. are often generalized to "Desulfovibrio" they are in fact discrete species with diverse phenotypes. Thus, the single sensitive host may not reflect a narrow host range for these phages, but rather a reflection of the limited panel of hosts, particularly of other *D. desulfuricans* isolates. Morphologically each phage was shown by TEM to have either siphophage morphology with long flexible non-contractile tails or podophage morphology with short non-contractile tails (**Figure 25**).

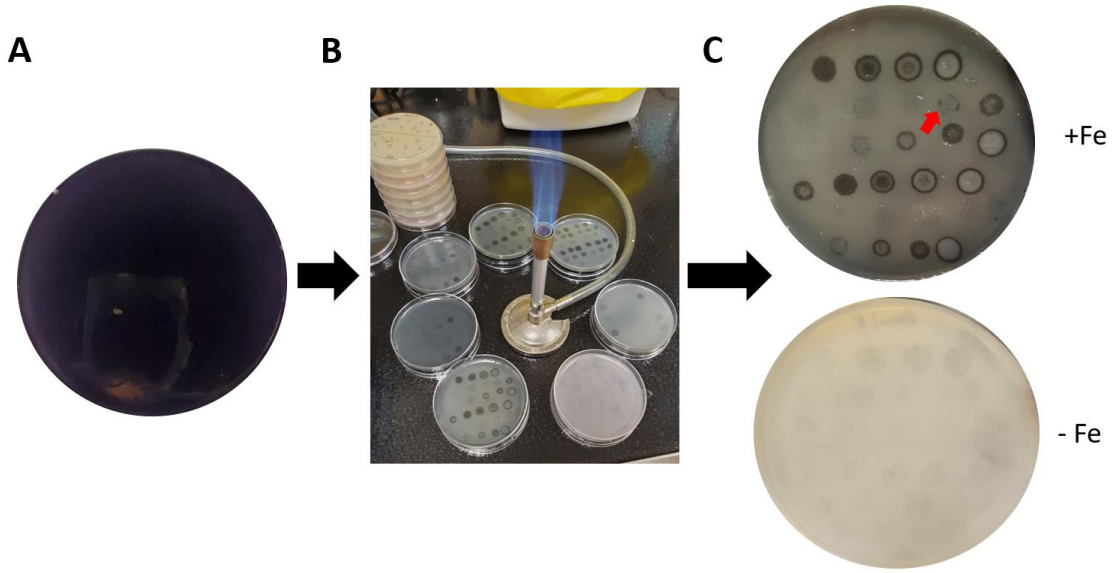


Figure 23. Demonstration of Method for Plaque Formation on *Desulfovibrio* Lawns. A). Visualization of *Desulfovibrio* lawn on MBfP plate when plated as indicated in methods. Note, the completely black plate with no visualization of plaques possible. B.) Plates containing a series of phage dilutions after O/N incubation, plates left out to incubate at RT aerobically for ~6 hrs. FeS begins to dissipate and plaques can be visualized. C.) Duplicate MBfP plates with and without Fe^{2+} to demonstrate enhanced plaque formation and visualization. Red arrow indicates individual plaques. Without Fe^{2+}, only opaque burn through could be seen, and individual plaques, if present at all, were extremely difficult to visualize.

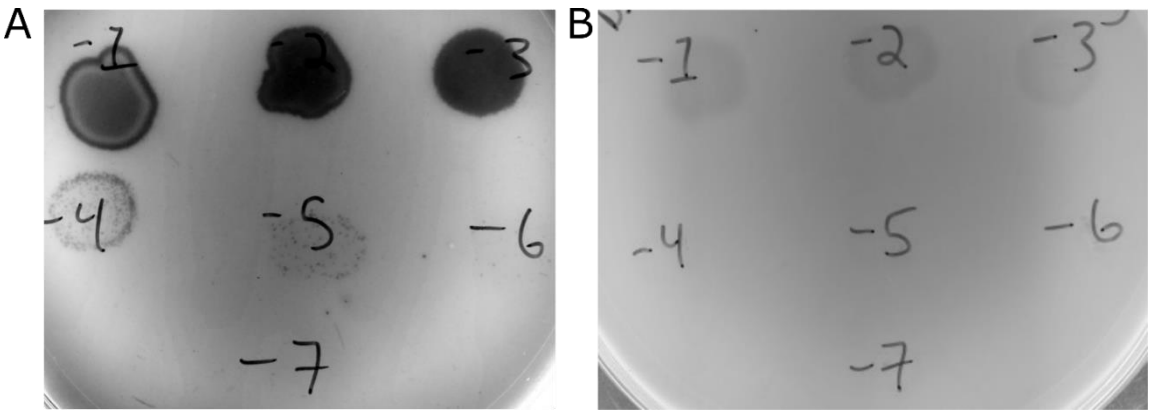


Figure 24. Further Demonstration of Phage Plaque Formation on *Desulfovibrio* Lawns Enhanced by Fe^{2+}.

A) 10-fold serial dilution of phage ddEPA on MBfP plates supplemented with Fe^{2+}. B) Replicate Plate without addition of Fe^{2+}. Image displayed in greyscale to enhance visualization of small plaques.

Phage	Morphology	Genome Size	Putative Lifestyle	Isolation Source	Host
ddEPA	Siphophage	54,177 bp	Temperate	Waste water	*D. desulfuricans* Edelweiss
dd4L	Siphophage	51,636 bp	Temperate	Waste water	*D. desulfuricans* Edelweiss
dd1	Podophage	39,698 bp	Virulent	Waste Water Influent	*D. desulfuricans* Edelweiss
dd15	Siphophage	43,383 bp	Temperate	Freshwater stream	*D. desulfuricans* Edelweiss
dd7442	Siphophage	54,745 bp	Temperate	Waste water influent	*D. desulfuricans* Edelweiss
dd41	Siphophage	55,020 bp	Temperate	Waste water influent	*D. desulfuricans* Edelweiss

Table 13. Summary of *Desulfovibrio* Phages Isolated.

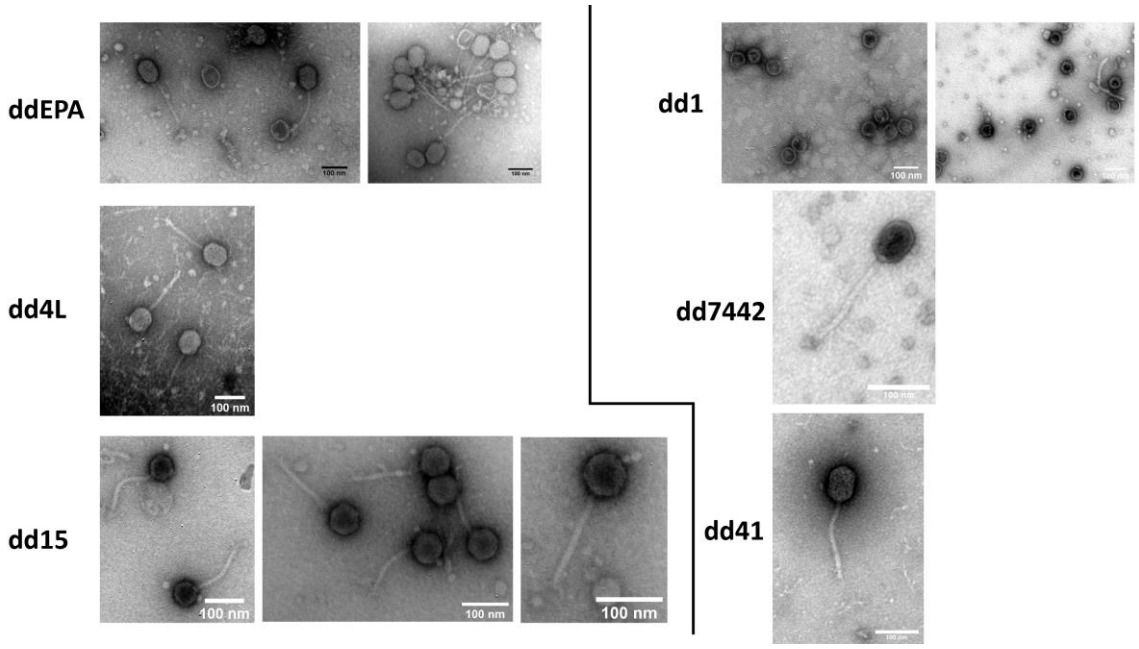


Figure 25. Desulfovibrio Phage Morphology.
Transmission Electron Micrographs of phage morphology. Stained with 2% uranyl acetate and imaged on a JEOL 1200 EX.

Desulfovibrio Phage Genomes

In general, each phage was shown to be highly novel by BLASTn, as can be seen in **Table 14**, with the closest genome matches deposited within Genbank sharing at most 34% coverage. Furthermore, these matches are partial metagenomic assemblies deposited as part of a recent large scale metagenomic phage assembly effort (Tisza & Buck, 2021). Other than these metagenomic sequences, coverage and similarity drops precipitously primarily to putative attB sites, the site of phage integration into the host

genome, and/or tRNAs. Despite the isolated phages infecting an environmental isolate of *D. desulfuricans*, the presence of related partial phage assemblies from human metagenomic sequencing and *D. desulfuricans* also being associated with the human gastrointestinal tract (Nakao et al., 2009) supports the relevance of the isolated phage's potential for applied uses to control human *Desulfovibrio* isolates. The exception being phage dd15, which contained no significant sequence homology to any deposited phage sequences, metagenomic or otherwise. Additionally, despite five of the six isolated phages being predicted to be temperate based on gene content, there were no significant matches to any prophages residing within deposited *Desulfovibrio* bacterial genomes. Relative to the other isolated *Desulfovibrio* phages, phage ddEPA, dd7442, and dd41 share the greatest similarity ranging from 66-75% nucleotide identity to one another (**Table 15**), while dd4L is the least similar while still belonging to this group of related phages sharing ~50% identity with ddEPA, dd7442, and dd41 (**Table 15**). In addition, each of these phages, ddEPA, dd7442, dd41, and dd4L share the same putative att site within host tRNA-Phe. Phages dd1 and dd15 were the most unique phages isolated sharing no more than 2% nucleotide identity with any isolated phages. Sequencing of the isolated phage also revealed the presence of a functional prophage element, ProddE, further discussed in the Chapter 5. However, speculatively, ProddE seemed to be induced in response to infection by specific isolated phages. ProddE was induced to levels allowing its assembly as a full length contig from assemblies of phage dd1 and dd41, as indicated by **Table 14**. PhageTerm (Garneau, Depardieu, Fortier, Bikard, & Monot, 2017) in this instance was widely inconclusive in determining *Desulfovibrio*

phage termini as well as packaging types. Given the extremely limited nucleotide identity to deposited bacterial and phage genomes, isolated *Desulfovibrio* phages were also screened on a protein network clustering basis using vContact2 (Jang et al., 2019) against the NCBI Bacterial and Archaeal Viral RefSeq v88. Consistent with nucleotide similarities, dd41, dd4L, dd7442, and ddEPA formed a high confidence single viral cluster with no other members of RefSeq (**Table 16**). Phage dd1 formed a viral cluster with a single phage within RefSeq, *Pelagibacter* phage HTVC010P. Also consistent with phage dd15's lack of any significant nucleotide similarity to deposited sequences, dd15 clustered with no phages within RefSeq v88 and was classified as an outlier. Analysis of the host genome showed the presence of a number of CRISPR protospacers targeting the phages isolated here (**Table 17**), or similar phages. The carriage of multiple near perfect CRISPR spacers for this group of phages implies frequent encounters between this *Desulfovibrio* host and related phage species. However, seemingly minor nucleotide alterations are sufficient for phage to escape CRISPR cleavage in this instance. BLASTn of the 42 spacer regions to the nr database only yielded hits to 8 of the 42 spacers in other *Desulfovibrio* genomes or metagenomic phage assemblies, suggesting the diversity of *Desulfovibrio* phage or other mobile genetic elements remains vastly undiscovered. Together, these results indicate the isolated phages are highly novel, and the phages isolated here represent three distinct classes of *Desulfovibrio* phage. These groups of phages will be individually discussed below.

Phage	Closest match in GenBank	Query Coverage	Percent Identity	Source	Accession	Putative Induction of ProddE
ddEPA	TPA_asm: Bacteriophage sp. isolate ctnGe20, partial genome	34%	90.38%	Human Metagenome	BK048101.1	
dd4L	TPA_asm: Bacteriophage sp. isolate ctnGe20, partial genome	30%	89.85%	Human Metagenome	BK048101.1	
dd1	TPA: Podoviridae sp. isolate ctMn125, partial genome	10%	67.32%	Human Metagenome	BK017135.1	X
dd15	n/a	n/a	n/a	n/a	n/a	
dd7442	TPA_asm: Bacteriophage sp. isolate ctnGe20, partial genome	31%	90.12%	Human Metagenome	BK048101.1	
dd41	TPA_asm: Bacteriophage sp. isolate ctnGe20, partial genome	33%	90.35%	Human Metagenome	BK048101.1	X

Table 14. Nucleotide Similarity of Isolated Phage to Existing Deposited Genbank Records.
Top hit reported as indicated by BLASTn against nr database with default parameters.

Putative induction of prophage ProddE also indicated as determined by co-assembly of

ProddE contig with the sequenced phage.

	ddEPA	dd4L	dd1	dd15	dd7442	dd41
ddEPA	100	45.92	1.76	1.32	63.86	74.65
dd4L	45.92	100	0.91	0.9	49.09	49.41
dd1	1.76	0.91	100	2.36	1.48	1.82
dd15	1.32	0.9	2.36	100	1.37	1.36
dd7442	63.86	49.09	1.48	1.37	100	66.36
dd41	74.65	49.41	1.82	1.36	66.36	100

Table 15. Percent Nucleotide Identity of Isolated Phages.
Percent nucleotide identity (Dice) as determined by ProgressiveMauve**.**

Phage	Cluster Status	Viral Cluster	Quality
Pelagibacter~phage~HTVC010P	Clustered	VC_196_0	0.1354
dd1	Clustered	VC_196_0	0.1354
dd41	Clustered	VC_244_0	0.9619
dd4L	Clustered	VC_244_0	0.9619
dd7442	Clustered	VC_244_0	0.9619
ddEPA	Clustered	VC_244_0	0.9619
dd15	Outlier	-	-

Table 16. Viral Clusters as Indicated by vContact2.
Isolated *Desulfovibrio* phage genomes were used as input and compared to NCBI Bacterial and Archaeal Viral RefSeq v88 using the default settings with BLASTp. Viral clustering supports 3 distinct groups of phages.

	Putative att site	# of Partial Crispr Protospacer Matches	Specific Protospacer #
ddEPA	tRNA-Phe	2	#5, #22
dd4L	tRNA-Phe	2	#14, #34
dd1	n/a	-	-
dd15	tRNA-Thr	-	-
dd7442	tRNA-Phe	3	#14, #22, #34
dd41	tRNA-Phe	2	#14, #22

Table 17. Phage-Host Interactions Inferred by Nucleotide Comparison.
CRISPRFINDER tool (Grissa, Vergnaud, & Pourcel, 2007) as well as BLASTn used to identify putative att sites as well CRISPR spacers targeting the isolated phage within the host.

Phage dd1

The podophage dd1 has the smallest genome of the phages isolated in this work with a genome of 39,698 bp. A genome map in comparison to its metagenome

assembled hit, ctMn125, can be seen in **Figure 26A**. Genomic content implies a virulent lifestyle as an integrase, partitioning system, as well phage repressors/antirepressors could not be identified. This makes dd1 the only virulent *Desulfovibrio* phage isolated in this study. Recognizable protein content of dd1 is relatively sparse, with few accessory genes beyond proteins recognizable proteins for DNA replication and virion formation. Phage dd1 does however encode a Cas4 CRISPR associated like-protein (domain: cl00641). Cas4 has recently been shown to be involved in spacer acquisition through Cas4 cleavage upstream of PAM sequences (Lee, Dhingra, & Sashital, 2019), perhaps implying host genome degradation as part of host take over or inhibition of host spacer acquisition. Furthermore, dd1 encodes a cytidine deaminase and an adjacent thymidylate synthase implying alteration of host nucleotide pools toward dTMP. Genomic arrangement of dd1 is split between a DNA replication locus and structural/lysis locus encoded in the opposite orientation. PhageTerm was inconclusive regarding packaging method and genomic termini, though direct read mapping via BowTie2 (Langmead & Salzberg, 2012b) suggests direct terminal repeats (DTR) as read distribution shows significant spike in coverage ~3x greater at putative the ~1.5kb DTR relative to the rest of the genome, similar to the sequence coverage distribution of phage T7 which contains DTR (Garneau et al., 2017).

Phage dd15

The siphophage dd15 has a genome of 43,383 bp, its genomic map that can be seen in **Figure 26B**. Consistent with dd15's lack of nucleotide similarity or protein clustering to other phages, a large portion of the genes of dd15 are either novel or

conserved hypothetical proteins. Like the majority of *Desulfovibiro* phages described here, dd15 is predicted to be a temperate phage due to the presence of an integrase (tyrosine recombinase). dd15 contains a putative att site within host tRNA-Thr for lysogen formation. Outside of the standard array of recognizable structural and replicative machinery for siphophage formation, dd15 also encodes a dCMP deaminase, implying an alteration of host nucleotide pools and a putative XRE family phage repressor. Phage dd15 also encodes an integration host factor (IHF)-like protein, speculatively implying alterations to host cell transcription.

Phage ddEPA, dd4L, dd7442, and dd41

The prolate siphophage ddEPA has 54,177 bp genome, a genome map can be seen in **Figure 27.** It should be noted that assembly of the ddEPA genome was difficult for SPAdes assembly program frequently resulting in small 10-15kb contigs, only an increase of the k-mer size allowed for the eventual assembly of full-length phage, or the use of metaSPAdes (Nurk, Meleshko, Korobeynikov, & Pevzner, 2017). As predicted by PhageTerm, ddEPA is thought to be permuted with redundant ends and follow headful (pac) packaging. PhageTerm was inconclusive on all other phages of this type, however given their nucleotide similarity, all phages of this type are opened accordingly to be syntetic with ddEPA. Outside of the standard siphophage structural proteins and those involved in DNA replication, ddEPA also contains a number of additional genes of note. ddEPA contains a tyrosine recombinase indicating a temperate phage lifestyle. No obvious method of lysogen repression could be identified; however, ddEPA does contain a recognizable anti-repressor displaying conserved domains of the Bro-N/COG3617

superfamily. ddEPA also encodes a Lar-like restriction alleviation protein, as well as a number of nucleases: a putative type III restriction family endonuclease, a cas9 family endonuclease, as well as a McrA family endonuclease, associated with the cleavage of modified DNA (Ramalingam, Prasad, Shivapriya, & Dharmalingam, 2007). Sequencing of the *D. desulfuricans* host and ddEPA, revealed ddEPA also possessed the appropriate attP site for phage integration within host tRNA-Phe. The *D. desulfuricans* host is predicted to have an active also contains two partial protospacer matches to ddEPA with 89% and 98% nucleotide identity to the ~35bp spacer regions, **Table 17**.

The siphophage dd4L has a 51,636 bp genome; its genomic arrangement is shown in **Figure 27**. Phage dd4L contains a tyrosine recombinase type integrase indicating a temperate lifestyle. In addition, dd4L harbors an identical attP site to phage ddEPA within a tRNA-Phe. The host harbors two partial protospacers with 97% and 88% nt identity to dd4L, these are distinct protospacers to those matching ddEPA, as can be seen in **Table 17**. dd4L contains a recognizable CI-like phage repressor of the LexA superfamily with both helix-turn-helix DNA binding domain (pfam13443) and catalytic peptidase site (cd0652). This repressor is absent from ddEPA imply alternative methods of phage repression in ddEPA. In addition, dd4L encodes a recognizable super infection exclusion protein, of the Sipho GP157 superfamily (cl05239).

The siphophage dd7442 has a genome of 54,745 bp, genomic map can be seen in **Figure 27**. Phage dd7442 suffered similar assembly difficulties to ddEPA, in the form of the assembly breaking into smaller contigs that can often be indicative of repetitive or tandem repeat sequences. These assembly issues were limited to the phages dd7442 and

ddEPA, suggesting that perhaps assembly difficulties may be a characteristic innate to this group of phages. Although dd4L and dd41, whom are also similar to dd7442 and ddEPA, suffered no assembly problems. Like dd4L and ddEPA, dd7442 harbors the same putative att site within the host tRNA-Phe, as well as a tyrosine-recombinase. Like dd4L, dd7442 encodes a CI-like family phage repressor for lysogen maintenance. Notably, immediately upstream of the CI-like gene start coordinate there is also encoded a XRE-family transcriptional regulator in the opposite orientation on the minus strand, implying a likely region for a lysogeny control switch. In addition to these putative lysis/lysogeny decision genes, dd7442 also contains a recognizable P22-like anti-repressor at a different locus embedded within the locus for structural proteins. The host contains three partial CRISPR spacer matches to dd7442 with 97%, 89%, and 88% nucleotide identity, **Table 17**.

The phage dd41 has genome of 55,020 bp, its genomic map can be seen in **Figure 27**. Like dd4L, dd7442, and ddEPA, dd41 encodes tyrosine recombinase, implying a temperate lifecycle, as well as a CI-like repressor and identical putative att site within host tRNA-Phe. Additionally, dd41 also carriers the P22-like anti-repressor, similarly to dd7442, embedded in the structural locus. Phage dd41 also encodes a ParB-like partitioning protein, however it lacks any recognizable accompanying ParA-like homolog. The host contains two partial CRISPR spacer matches to dd41 sharing 94% and 89% nucleotide identity (**Table 17**).

In summary, ddEPA, dd41, dd4L and dd7442 represent a similar but novel group of temperate phages infecting *Desulfovibrio* spp. characterized by siphophage

morphology, sharing ~50-70% genomic content across ~55 kb genomes, and a common att site within tRNA-Phe. This high conservation of nucleotide identity across >50% of the genome likely explains the assemblage of portions of the genome from human metagenomic sequences, as indicated by Table 14. Therefore, the conserved stretches of sequences in this phage group are likely overrepresented amongst *Desulfovibrio* phages in metagenomic sequencing pools. The metagenomic element ctnGe20 is included in the genomic comparison map to demonstrate limitations of metagenome phage assemblies, as only approximately half of the genome is captured. The comparative genomic map of **Figure 27**, demonstrates a conserved genomic arrangement. While many of the genes remain hypothetical, the genomes appear to be broadly split between three major loci. Firstly (I), a structural locus in the positive sense orientation that appears the most conserved between the phages was identified containing a TerL, TerS, major and minor capsid proteins, as well as tail fiber, and tape measure components. Next (II), this phage group contains a locus in the minus orientation, containing a conserved integrase as well as variable presence of other lysogeny associated proteins as described above. This locus also contains a conserved ERF-like (Essential Recombination Function) protein. Notably, this region also encodes are large putative DNA Methylase protein with conserved domains for helicase activity (cd17919), helicase ATP binding site (cd18793), and DNA methylation (pfam01555). This DNA methylase is notably absent in phage dd4L, and in ddEPA and dd7442, this gene contains an intein/homing endonuclease. In the coding regions downstream of the ERF-like protein, individual gene content becomes highly variable before returning to a more highly conserved locus (III)

containing the putative suite of lysis proteins (endolysin, holin, and spanin proteins). Relatively few proteins related to DNA replication could be identified in this group of phages implying cryptic replication machinery or heavy reliance on host proteins for replication. Interestingly, as mentioned previously, although not universal, half of the phages of this group described here suffered from genome assembly issues (ddEPA and dd7442), further discussed below, which may reflect a currently unknown aspect these phages' biology which confounds de Bruijn graph assemblers, however the isolation of additional members of this group of phages is necessary to confirm this trend.

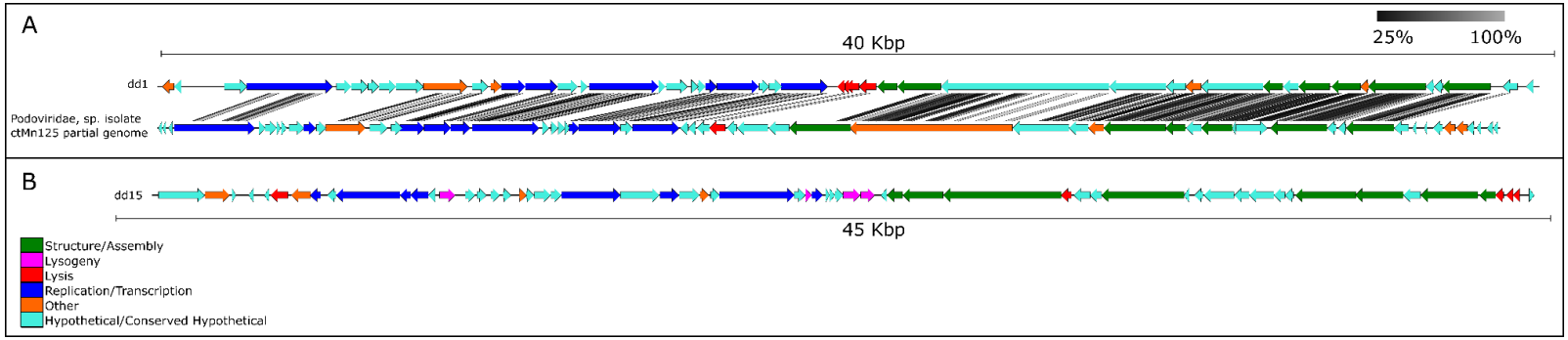


Figure 26. Genomic Maps of Phage dd1 and dd15.
A) Phage genomic map of phage dd1 and its comparison to the metagenome assembly of the only similar phage ctMn125, partial genome. Connecting lines indicate amino acid similarity via tBLASTx and generated with EasyFig using 25% similarity and a minimum length to be displayed of 50. B) Genomic map of dd15. A and B genes color coded according the figure legend with putative functions.

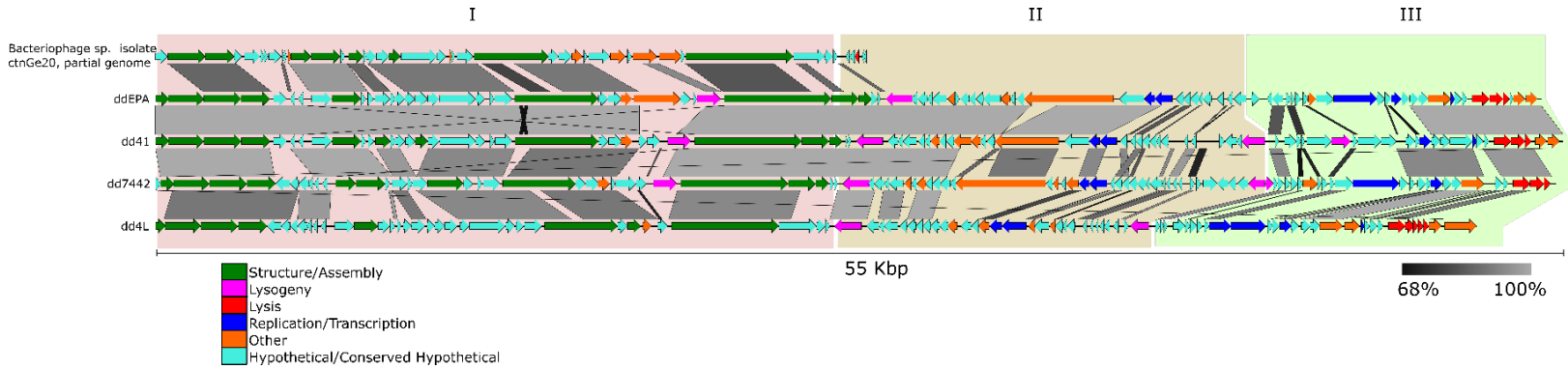


Figure 27. Genomic Comparison of ddEPA-like Phages.
Similar phages ddEPA, dd41, dd7442, and dd4L, as well as the metagenomic assembly ctnGe20, compared via BLASTn and visualized via EasyFig with minimum length of a hit to be visualize 65bp and minimum E-value .001.

Confirmation of Phage ddEPA Lysogeny and Insights into Genomic Assembly Errors

Given phage ddEPA was the first phage isolated, it received additional levels of characterization. Isolation of phage ddEPA resistant colonies were shown to be lysogens through spontaneous induction of ddEPA, as indicated by plaque formation by ddEPA lysogen supernatants. Sequencing of three putative lysogens also showed assembly of either approximately full length ddEPA contigs or ddEPA split between a number of contigs, mirroring the initial assembly problems of purified phage ddEPA DNA. Surprisingly, the lysogenic version of ddEPA was not a complete match for the version assembled from purified phage DNA. For simplicity's sake, the purified phage assembly of ddEPA will be referred to as ddEPA-P and the lysogen assembly as ddEPA-L. ddEPA-L shared only 73% query coverage and >99% nucleotide identity with the ddEPA-P version displaying gaps in alignment. The sequences of the two assembled lysogenic versions of ddEPA-L also displayed 96% alignment coverage relative to one another, but contained scattered alignment gaps of ~100-600 bp across the genome despite high read coverage. Nor was this ddEPA-L conformation a match any of the other isolated phages. The ddEPA-L assembly also contained novel gene insertions not present in the ddEPA-P assembly such as an HNH endonuclease and an alternate major capsid protein. To assess which version the ddEPA-P or ddEPA-L was the "true" genomic conformation, discriminatory primers designed to give differential product sizes across gapped regions of the genome were constructed and utilized on the purified phage ddEPA-P DNA and DNA from the lysogen. These results showed that both

genome conformations, ddEPA-P and ddEPA-L, were present in purified phage DNA preparation as indicated by **Figure 28**. However, the ddEPA-P genomic conformation was seemingly lost within the lysogen as indicated by PCR of DNA extracted of the lysogen. Notably, the spontaneous plaques originating from the ddEPA-L lysogen supernatant were not able to be subcultured despite repeated attempts indicating a reduced viability. BowTie mapping of the short sequencing reads from the non-lysogenized host and ddEPA-P against the novel elements of ddEPA-L showed their presence indeed originated from the ddEPA-P DNA prep and not the host. However, the elements apparently absent from ddEPA-L assembly notably included the integrase as well as the ERF superfamily recombination protein amongst others; these elements were confirmed absent by read mapping from the ddEPA-L short sequencing reads. This evidence shows presence of both genomic assembly conformations present in ddEPA preparations, though the reason for their presence remains unclear. It could perhaps be due to a simple mixed phage population due to a contaminating similar phage (although why they were not capable of separation by multiple rounds of subculture from single plaques remains unclear), frequent recombination events during infection leading to the presence of multiple genome confirmations, the presence of a phage parasitic element, a dependent phage necessitating a helper phage, perhaps indicated by the inability to subculture plaques originating from the lysogen, or some other heretofore unknown aspect of this class of phages. These observations are reported here in an effort to aid any future sequencing of *Desulfovibrio* phages belonging to this group of phages that may display similar assembly issues.

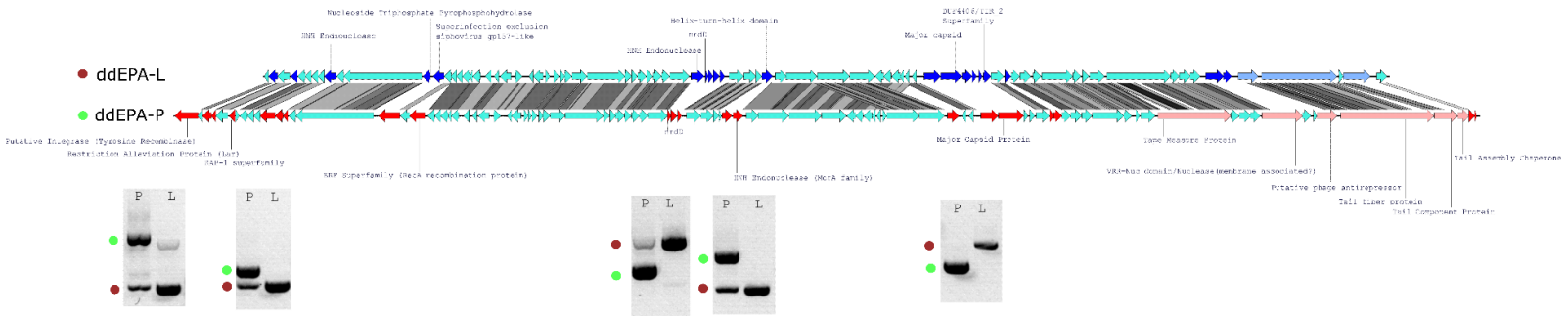


Figure 28. Comparison of Lysogenic and Purified Phage ddEPA DNA assemblies.
Genomic comparison of ddEPA-L assembly vs the ddEPA-L assembly from sequencing of the lysogen. EasyFig comparison via tBLASTx. Red labeled genes indicate genes absent from the lysogen assembly but present in phage assembly. Blue indicated regions present in lysogen assembly but absent in phage assembly. Light pink and light blue indicate genes disrupted but sharing homology. Bottom, gel insets display representative PCR confirming presence of both genome conformations at the region above the inset. Green dot indicates expected PCR product from the purified phage DNA assembly and maroon dot indicates expected product size from the lysogen. Lane P indicates phage DNA as template, lane L indicates lysogen DNA as template.

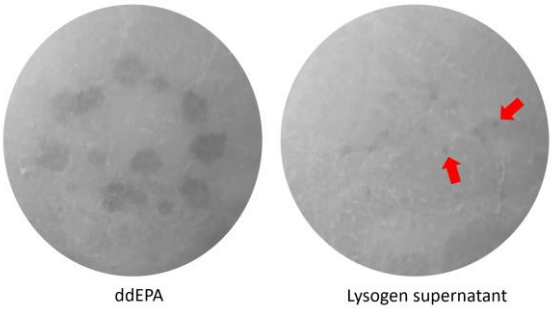


Figure 29. Plaque Formation From ddEPA Lysogen Supernatant.
Comparison in size of ddEPA plaques from ddEPA phage stock vs the filter sterilized lysogen supernatant plated on WT (non-lysogen) host. Note plaques from the lysogen supernatant indicated are exceptionally small relative to ddEPA denoted by the red arrows. Plaques from the lysogen supernatant were incapable of being propagated.

Putative Receptor of Phage ddEPA

In addition to the ddEPA lysogens sequenced, a number of non-lysogen phage resistant mutants to ddEPA were isolated and sequenced. On these mutants, ddEPA displayed and efficiency of plating (EOP) of ~10^{-3} relative to the wild type host. Mapping of sequencing reads of these resistance mutants to the WT host showed a number of indels or snp mutations localized to the apparent LPS synthesis locus. Examination of these mutations and their impact on the open reading frame (ORF) in which they reside showed an early -1 frameshift in a putative Wzt homolog (cd03220 and cd10147) (HHpred: 7K2T), a premature stop codon in a protein containing a conserved domain of unknown function (DUF6942) with predicted structural similarity to a glycosyl transferase (HHpred: 1G9R_A), and a premature stop codon in a glycosyl

transferase family 10 protein. A summary of these mutations and schematic of their locations can be seen in **Table 18** and **Figure 30**. The KpsT/Wzt ABC transporter subfamilies are involved in export of polysaccharides such as LPS, and Wzt in particular is associated with transport of O-antigen (Caffalette & Zimmer, 2021; Cuthbertson, Powers, & Whitfield, 2005; Samuel & Reeves, 2003). The premature stop codon of the conserved protein containing DUF6942 has no known function, although given its location in this locus and structural match to a glycosyl transferase, it likely also plays a role in LPS synthesis. While the snp identified in resistant mutant ddEPA-R4, leads to a premature amber mutation in the glycosyltransferase 10 family (GT10) protein within the same locus. None of these proteins are predicted to be outer membrane proteins, and thus would be inaccessible to the phage tail fibers, therefore the predicted product of their pathway, the generation of LPS (particularly O-antigen), is likely the phage receptor for ddEPA. However, given the reduction in EOP plating is not more severe, it is possible likely that there exists a second receptor for phage ddEPA and LPS plays a role in ddEPA adsorption but is not exclusively the receptor.

Resistant Mutant	Mutation	Type	Gene	EOP
D. desulfuricans Edelweiss ddEPA-RL	Ser32Opal	Nonsense	Wzt homolog	~10^{-3}
D. desulfuricans Edelweiss ddEPA-R1	28delC	-1 Frameshift	DUF6942 protein (Putative Glycosyl transferase)	~10^{-3}
D. desulfuricans Edelweiss ddEPA-R4	Gln46Amber	Nonsense	Glycosyl transferase 10	~10^{-3}

Table 18. SNP/Indel Analysis of ddEPA Resistant Mutants Suggest LPS Plays a Role in Phage Adsorption.

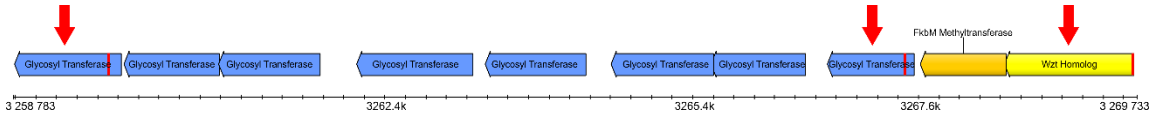


Figure 30. Localization of SNPs/Indels to the Putative LPS Synthesis/Export locus of *D. desulfuricans* Edelweiss.
Mutations conferring ddEPA resistance localized to the putative LPS (O-antigen) synthesis locus. Blue indicate glycosyl transferase genes. Orange indicates putative FkbM Methyltransferase. Yellow indicates Wzt homolog. Red arrows indicate genes containing mutations described in Table 18. Red line indicates mutated position in the ORF.

Single Cell Demonstrations of *Desulfovibrio* Lysis

Given that the evidence of phage lytic activity in *Desulfovibrio* is extremely limited and often restricted only to the identification of virus like particles upon UV or mitomycin C induction (Handley et al., 1973; Seyedirashti, Wood, & Akagi, 1991b; Seyedirashti et al., 1992), we sought to evaluate the lytic ability of isolated *Desulfovibrio* phages to assess the plausibility of their use for applied purposes. Some phages displayed potent lytic ability to clear a liquid culture of *Desulfovibrio* leading to formation of clumped aggregates, presumably cell debris or formation of precipitates from the medium as can be seen in **Figure 31**, while others phages resulted in a plateau of the culture optical density. Furthermore, as measured by formation of FeS, *Desulfovibrio* phage ddEPA was capable of inhibiting growth of the host for up to 3 days at higher phage concentrations, **Figure 31**. Select phages were examined at the single

cell level to observe individual lysis events. Representative time lapses of individual lysis events for phages ddEPA, dd4L, and dd15 can be seen in **Figure 32**. It should be noted that observed lysis events may be subject to artifacts due to brief oxygen exposure, however no such lysis events were observed in no phage control samples. In general, phages displayed classic lysis phenotypes marked by formation of a membrane blebs/bulging prior to explosive cell lysis marked by loss of cell integrity and a reduction in phase dark contrast. Formation of membrane blebs can be seen in **Figure 32,** as indicated with the red arrows. Additionally, dd4L also seemingly contains lytic proteins that retain activity well after the initial lysis event which continues to degrade the host membrane, a trait not observed in dd15 or ddEPA. While formation of membrane blebs was observed in ddEPA, no lysis events were observed during the recorded period, and only cells forming blebs or cells that had already formed blebs were observed. This observation is perhaps consistent with observed plateau of culture ODs rather than lysis. It is possible an additional mechanical force not present under these tested conditions is necessary for final lysis of these membrane blebs observed in ddEPA, or this lack of observed cell lysis is an artifact due to host oxygen exposure. Finally, it was also common for the release of a phase dark pellet after lysis, indicated by the blue arrows. These are presumed to be akin to iron sulfur bodies (ISB) for sequestration of iron (Toso, Javed, Czornyj, Gunsalus, & Zhou, 2016), although ISBs have yet to be described in *Desulfovibrio*. These results demonstrate that, as expected, *Desulfovibrio* phage are capable of efficient host cell lysis.

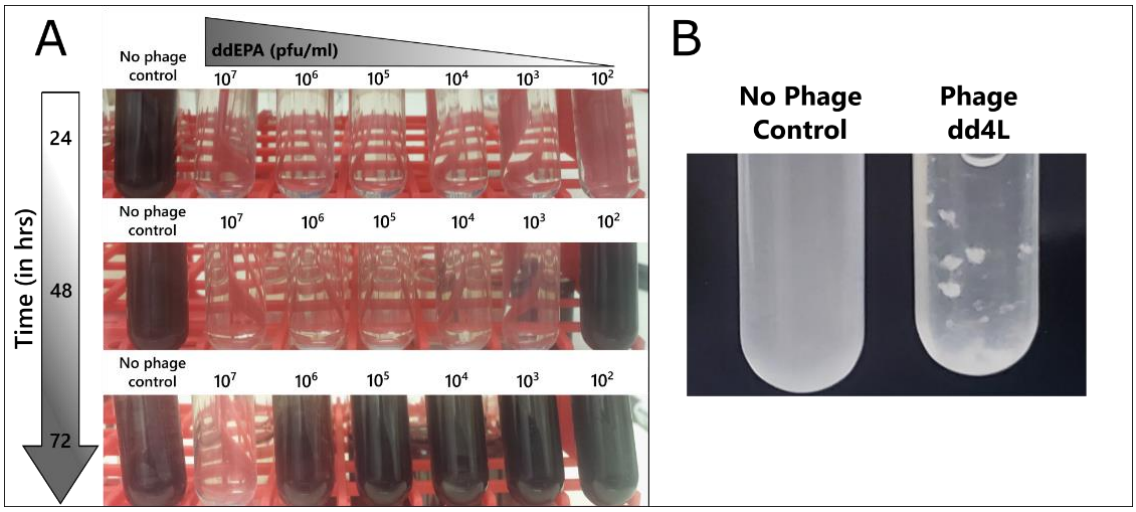


Figure 31. Examples of *Desulfovibrio* Inhibition of Growth and Lysis.
A.) MBfP supplemented with Fe was inoculated with 1/10th volume of *Desulfovibrio* culture and inoculated with serial phage dilutions of ddEPA to concentration final concentrations ~10^7 to 10^2 pfu/ml, and monitored for 3 days FeS formation as an indicator of H2S production and *Desulfovibrio* growth. B.) Demonstration of lytic capabilities of phage dd4L. Rapidly growing culture of *Desulfovibrio* at grown to an OD of ~0.3 was inoculated with phage dd4L to an MOI of ~3 and following O/N incubation at 37 °C, cultures had lysed displaying clumped aggregates that were commonly seen post lysis.

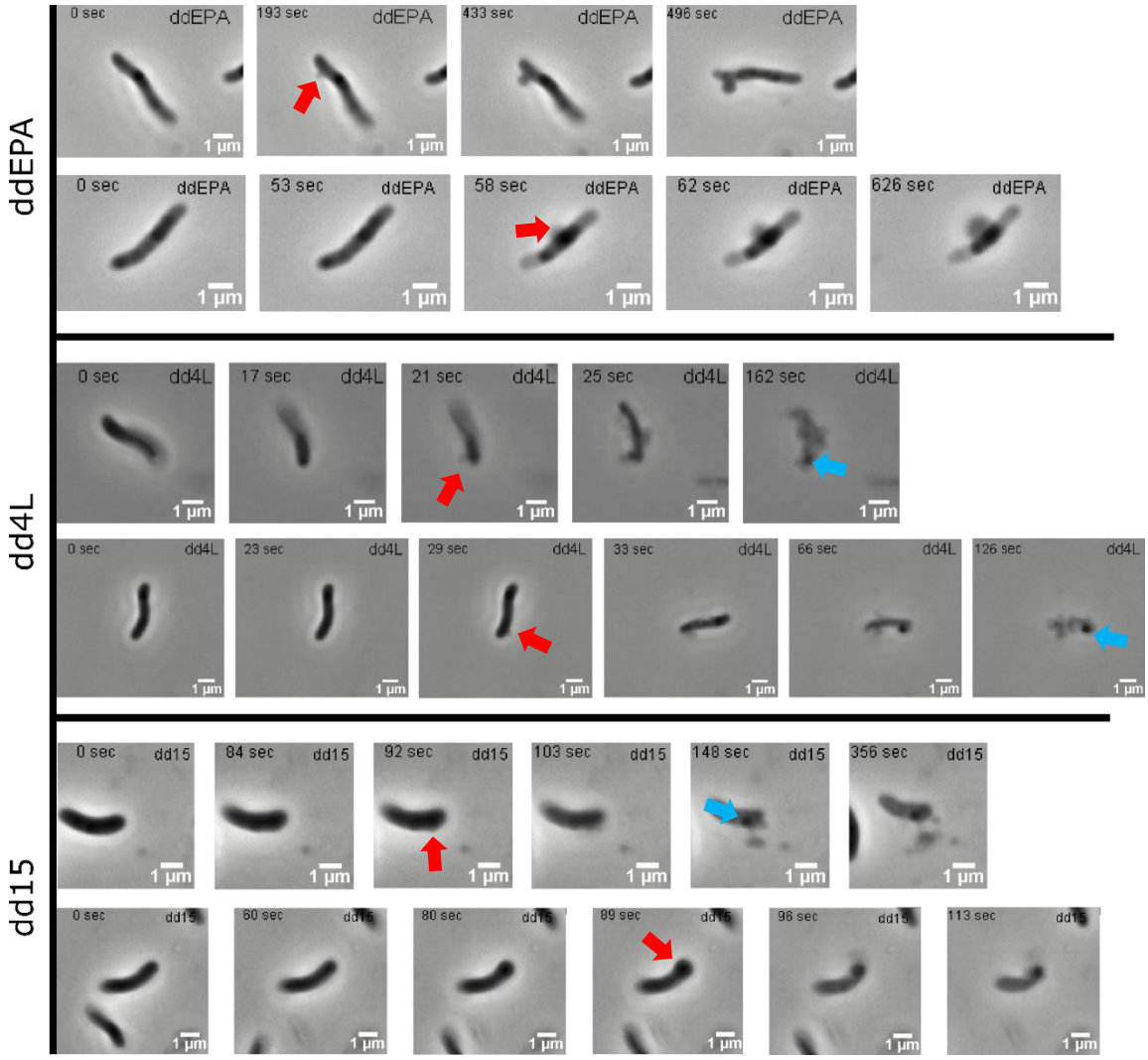


Figure 32. Representative Single Cell Lysis Events of Select Phages.
To observe individual lytic events of select phages, rapidly growing *Desulfovibrio* host was infected at an MOI ~1 and incubated for 2 h, the time previously determined time at which noticeable lysis begins to occur or a plateau of culture OD. After 2 h, wet mount slides with coverslip were prepared within the anaerobic chamber and rapidly move to microscope and imaged with time lapse phase microscopy for a 20 min period. Red arrows indicate initial formation of membrane blebs/bulging preceding lysis. Blue arrows indicate a phase dark pellet frequently deposited post lysis, presumed to be iron

sulfur bodies. Time stamps are relative surrounding observed lysis events and do not indicate time post infection.

Inhibition of *Desulfovibrio* Growth

To further evaluate and demonstrate the potential of *Desulfovibrio* phage for applied uses to reduce/control *Desulfovibrio* growth, each phage's ability to inhibit *Desulfovibrio* growth was assayed. In brief, MBfP media was inoculated with $1/10^{th}$ volume of growing *Desulfovibrio* culture and infected with the described isolated phage to a multiplicity of infection (MOI) of ~0.2 in a Hungate tube. Tubes were incubated at 37 °C with shaking and the OD_{600} was monitored for 6 days. As can be seen in **Figure 33**, most phages, with the exception of dd1 and dd41, were capable of inhibiting growth of *Desulfovibrio* for the duration of the experiment never approaching the optical density of the no phage control (black) during the experiment, although the phage infected cultures are slowly trending toward an increase in OD_{600}. This slow rise in OD may represent lysogen formation or growth of resistant mutants, however it is at a much slower rate implying a fitness cost or continued pressure by the phage. Notably, phage dd1 (red), the only virulent phage isolated, had the weakest ability to inhibit growth with negligible alterations to *Desulfovibrio* growth compared the control culture. The only other phage that had only a mild effect of growth of *Desulfovibrio* was phage dd41 (dark orange), which showed initial suppression of growth followed by rapid rise in OD. Notably, despite their lack of similarity, dd41 and dd1 were also the only phages whose infection suggested a high level of induction from resident prophage ProddE, **Table 14**.

Thus, the observed lack of growth inhibition may represent a pseudo-abortive infection in which dd1 or dd41 leads to induction of ProddE before dd1 and dd41 can finish their own infection cycle, leading to release of ProddE into a population of insensitive hosts and resulting in minimal reductions in OD by dd41 and dd1. The inhibitory abilities of the phage displayed here support *Desulfovibrio* phage may be suitable for targeted reductions of *Desulfovibrio* in environmental or intestinal contexts, and even single phages at relatively low MOIs are capable of inhibiting *Desulfovibrio* growth for sustained periods of time. However, despite displays of their lytic ability, the majority of phages isolated here are predicted to be temperate, so the use of such phages would require engineering or isolation of virulent mutants to convert them to a virulent lifestyle.

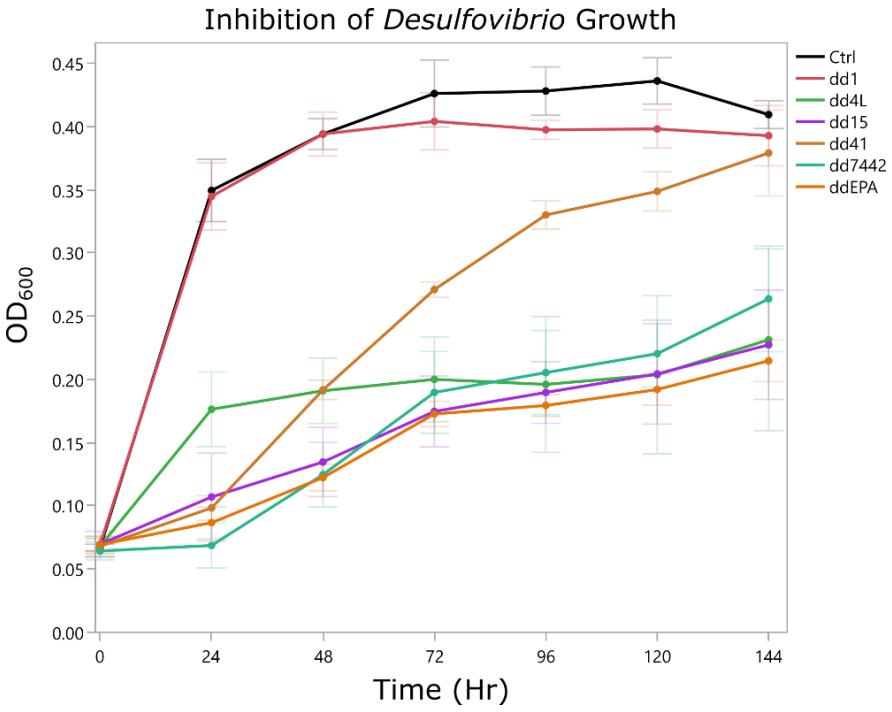


Figure 33. Demonstration of Phage Inhibition of *Desulfovibrio* Growth. MBfP media without Fe inoculated with *Desulfovibrio* culture and infected with the phage indicated in the figure legend to an MOI of 0.2. Cultures incubated at 37°C with agitation and growth monitored for a 6-day period. Error bars indicate SD.

4.3. Conclusion

In this work, we present an in-depth examination of *Desulfovibrio* phages isolated from the environment, report characteristics of their genomes, morphology, and displays of their lytic potential. To our knowledge this work represents the most thorough reporting of *Desulfovibrio* phages to date, along with a description of methods to facilitate future isolation of *Desulfovibrio* phage. Six phages in total were identified: ddEPA, dd41, dd4L, dd7442, dd15 and dd1 all of which are highly novel. ddEPA, dd41, dd4L and dd7442 represent a novel group of phages with little similarity to existing

phages both on nucleotide and a protein level. Phages dd1 and dd15 are unrelated to the other phages described in this work and each represents its own novel group of phages. Outside of metagenome assemblies, dd1 and dd15 have very little similarity to deposited phages. With the exception of dd1, which showed the poorest lytic performance in this study, all other phages isolated are predicted to be temperate phages. Despite their temperate nature, many show a strong inhibition of *Desulfovibrio* growth supporting their potential use in therapeutic applications, however this use would likely require modification to eliminate lysogen formation. While this modest study only reports six phages, the isolation of five out of six temperate phages may be indictive of temperate phages infecting *Desulfovibrio* being more common and makes isolation of strictly lytic phages more difficult, as has been reported for other anaerobes, like *C. difficile* (Hargreaves & Clokie, 2014). However, additional study of *Desulfovibrio* infecting phages is required to support dominance of temperate *Desulfovibrio* phage in the biosphere, further elucidate hypothetical gene content, and to validate a number of the trends reported here.

4.4. Materials and Methods

Bacterial Strains

For isolation of SRB species, Postgate Medium B (Postgate, 1984)(DSMZ_Medium63), as well as Modified Barr's Medium for Sulfate Reducers (ATCC Medium: 1249 Modified Barr's Medium For Sulfate Reducers), supplemented to a final concentration of 3.6mM Iron (II) sulfate Heptahydrate from a 0.22 μm PES filter

sterilized solution of 5% w/v Iron(II) sulfate heptahydrate solution when FeS indicator was desired and supplemented with resazurin sodium salt solution to final concentration of 995 nM. A minimum of 2 h prior to inoculation, reducing agent/oxygen scavenger in form of filtered sterilized 3% (w/v) L-Cysteine was added to broth media formulations to a final concentration of 3.8 mM. EDCM medium as described by Chen et al. was used for additional selection of *Desulfovibrio* species (Chen et al., 2019). To isolate *Desulfovibrio* species, ~1-2 g of sample, for an approximate 1/20 (w/v) dilution, was added to anaerobic Hungate tubes filled with reduced medium and supplemented with Fe. Hungate tubes were placed within COY brand vinyl anaerobic chamber with 5% H_2, 5%CO_2, and 90% N_2 atmosphere, incubated at 37 °C for a minimum of 14 days, and monitored for formation of FeS. Tubes positive for SRB growth were then diluted in 10-fold serial dilution and plated on SRB media supplemented with Fe^{2+} and amended with 1.5% Bacto agar. Plates were then incubated for a minimum 14-days and monitored for individual black colonies. Black colonies were then sub-cultured three times to ensure clonality and the 16S region amplified via colony PCR using Bact8F (Edwards, Rogall, Blöcker, Emde, & Böttger, 1989) and 1492R (Galkiewicz & Kellogg, 2008; Lopez et al., 2003) primers and sent for subsequent Sanger sequencing. Resulting 16S sequences were then classified to the genus or species level by BLASTn to the NCBI 16S ribosomal RNA sequence (Bacteria and Archaea) database as well as using the Ribosomal Database Project (RDP) classifier feature (Cole et al., 2014).

Phage Isolation and SRB Medium for Phage Isolation

All phage work was performed using a modified version of the Modified Barr's medium for sulfate reducers termed Modified Barr's medium for sulfate reducers for phage (MBfP). Alterations primarily consisted of removal of sodium citrate and calcium sulfate dihydrate to eliminate chelation of divalent cations and formation of precipitate often associated with Modified Barr's medium and Postgate Medium B, plus the addition of $CaCl_2$ solution post autoclaving to account for Ca^{2+} requirement. MBfP broth formulation 1 L was prepared in the following manner of Component 1: 2 g $MgSO_4$; 1 g NH_4Cl; 250 µl of .1% Resazurin solution (final concentration of 995 nM); 400 ml of H_2O; Component 2: 0.5 g K_2HPO_4, 200 ml of H_2O; Component 3: Sodium L-lactate 3.5 g, Yeast Extract 1 g, 400 ml of H_2O. Prepare each component separately, adjust each component to 7.5 pH. Autoclave. Post autoclave combine components aseptically. Add 5 ml of sterile 1 M $CaCl_2$ to final concentration of 5 mM, this will cause medium to become slightly cloudy. A minimum of 2 h prior to use aliquot necessary volume, and added 100 µl of filter sterilized 3% L-Cysteine per 5 ml under anaerobic conditions for a final concentration of 3.8 mM. After addition of L-Cysteine medium will become clear, and after ~2 h solution will become turn from pink to colorless as indicated by resazurin. For MBfP "t-top" for soft agar overlays: prepare as indicated above with the following modifications: To component 1, add 5 g Bacto agar and 0.1 g Thioglycollic acid. Prior to use, melt and aliquot 5ml molten t-top for soft agar overlays, add 100 µl of filter sterilized 3% L-Cysteine per 5 ml (final conc. 3.8 mM), and maintain aliquots under anaerobic conditions in 45-55 °C heat block. For agar plate formulations, components were prepared as described above for broth, with the following modifications for 1 L: To

component 1, add 1 g $CaSO_4$ x $2H_2O$; 15 g add Bacto agar; and 0.1 g Thioglycolic Acid. Omit $CaCl_2$ solution post autoclaving given presence of $CaSO_4$. Resulting solution contains small amount of precipitate; swirl to evenly distribute while pouring plates. When FeS was desired, 5% w/v Iron(II) sulfate heptahydrate solution was added to any above formulations to a final concentration of 3.6 mM (100 µl per 5ml media).

To enrich for *Desulfovibrio* phage, environmental samples were prepared as described in Chapter 3, and samples homogenized into MBfP under aerobic conditions. 10 ml of resulting sterile supernatants were added to 35 ml pre-reduced MBfP medium under anaerobic conditions and allowed to sit for 2 h minimum before addition of *Desulfovibrio* inoculum. A 10% (v/v) inoculum (5 ml) of *Desulfovibrio* inoculum from a relatively rapidly growing culture at an OD_{600} of 0.3 to 0.5 (a ~24 to 48 h culture) was added and enrichment cultures were grown anaerobically at 37 ˚C for 48-72 h. After 48-72 h, the resulting enrichment culture was spun at 4,750 rpm for 20 min at 4^{o}C and supernatant filter sterilized through a 0.22 µm PES filter. Sterilized supernatant was then screened for plaque formation and/or zones of clearing on soft-top agar overlays and inhibition of growth in liquid culture. For screening via top-agar overlays, in general 100 µl of O/N growth or a 100 µl of exponentially growing culture was found to be insufficient for generation of confluent lawns, instead concentrated cell pellet as had been described previously was utilized (Kamimura & Araki, 1989). 5 ml per plate of a rapidly growing culture between an OD_{600} of 0.3 - 0.6 was sealed tightly into 50 ml conical tube removed from chamber and briefly spun at 4,750 rpm for 5 min and rapidly returned to chamber. Under anaerobic conditions the supernatant was removed, save 500

µl per plate to resuspend the cell pellet. 500 µl of this concentrated cell pellet was added to 5 ml of MBfP top agar overlays as described above, and poured onto MBfP agar plates supplemented with Fe. Note, only the bottom 1.5% agar plate was supplemented with Fe in this method. To accommodate the 37 °C temperature of the anaerobic chamber in use, lawns were poured on a leveled steel plate held at -20 °C just prior to pouring of top-agar lawns. Without this chilled steel, plate top-agar lawns took upwards of 30 min to solidify resulting in malformed and inconsistent of bacterial lawns. Enrichment supernatant was then either spotted resultant top agar lawns or preabsorbed to host cells prior to inoculation to screen full plates for zones of clearing or plaques. After, O/N incubation the method described above should result in a plate completely black in appearance obscuring any ability to visualize plaques or burn through. Plates are removed from the chamber and allowed to incubate in an aerobic environment for (6-8 h) leading to the oxidation of FeS, eventually rendering the plates readable. It was found this method of inclusion of FeS in the agar plate greatly enhanced plaque formation and visualization, as the plaques retained their black coloring for a longer period of time relative to the rest of the plate (see **Figure 23**). Plaques isolated this manner were subcultured 3x to ensure a clonal isolate. To screen by liquid assay, 5 ml of pre-reduced MBfP were inoculated with $1/10^{th}$ volume of rapidly growing *Desulfovibrio* culture (0.3 - 0.5 OD) and 100 µl of phage enrichment supernatant added. Culture was monitored for inhibition of growth relative to a control via OD, and enrichments displaying inhibition of growth were filter sterilized, serially diluted, and reinoculated. This process was continued with the highest dilution of phage still inhibiting growth 3 times, or until

plaque purification 3x could be performed to ensure clonality. Note, all phages initially isolated via liquid propagation were plaque purified 3 times to ensure clonality.

Phage and Host Genomic DNA and Analysis

Resulting genomes of isolated phages and bacterial hosts were extracted, annotated and analyzed as described in Chapter 3. Identification of a putatively active CRISPR system and associated spacers in *D. desulfuricans* Edelweiss was determined by analysis of *D. desulfuricans* host using CRISPRFinder (Grissa et al., 2007).

Isolation of Putative Lysogen and Phage Resistant Mutants

Resistant mutants and/or lysogens of ddEPA were obtained via isolating bacterial survivors from zones of clearing due to high concentrations of phage. Resulting survivor colonies were subcultured 3x to ensure clonality. Efficiency of plating (EOP) of ddEPA on resulting mutants was performed as described in Chapter 3. Putative lysogens of ddEPA were screened for plaque formation, by filter sterilizing supernatant of a rapidly growing ddEPA lysogen culture and spotting it onto a lawn of WT non-lysogenized host. Bacterial Host DNA was isolated and analyzed for SNPs and indels as described in Chapter 3.

Demonstration of *Desulfovibrio* Lytic Activity

To observe individual lytic events of select phages, 100 µl of rapidly growing *Desulfovibrio* host was infected at an MOI ~1 and incubated for 2 hrs, the time previously determined time at which noticeable lysis begins to occur or a plateau of culture OD. After 2 h, wet mount slides with coverslip were prepared within the anaerobic chamber and rapidly moved to a Zeiss Axio Observer 7 microscope and

observed by time lapse phase microscopy with images taken every second for 20 min. Representative observed lysis events with notable snap shots shown in **Figure 32** display phage inhibition of growth as determined by FeS formation as a proxy for H_2S production, as in **Figure 31**. Culture tubes containing serially diluted phage concentrations were inoculated with ~1/10th volumed culture inoculum to a final volume of 5 ml, supplemented with 900 μM Iron(II) Sulfate Heptahydrate and monitored for 72 h for formation of FeS indication of *Desulfovibrio* growth. Reduction in Fe concentration was necessary to eliminate immediate formation of black precipitate due to highly reducing starting inoculum.

To display phage inhibition of *Desulfovibrio* growth and compare all phage isolated via OD, 5ml MBfP was inoculated with 1/10th volume of rapidly growing *Desulfovibrio* culture in hungate tubes and inoculated with 10^5 pfu/ml of the indicated phage to an MOI of ~.02. Starting inoculum determined by most probable number serial dilutions into MBfP medium supplemented with Fe to use FeS as indicator. Hungate tubes were sealed and incubated with agitation at 37 °C and OD monitored for 6 days. Mean of 3 biological replicates shown in Figure 33, error bars indicated SD.

5. SHEEP IN WOLVES' CLOTHING: TEMPERATE T7-LIKE BACTERIOPHAGE LINEAGES AND THE ORIGINS OF THE *AUTOGRAPHIVIRIDAE*

5.1. Introduction

Bacteriophage T7 is one of the most well-studied phages in biology, with its isolation traceable back to the early years of phage biology and a privileged status as one of the paradigm "seven dwarves" of *E. coli* phages (Abedon, 2000; Delbruck, 1946; Demerec & Fano, 1945). T7 is considered an archetype virulent phage due to its short latent period (as brief as ~15 min), catabolism of the host genome, relatively compact genome, minimal reliance on host proteins, ability to aggressively clear cultures, and formation of large clear plaques capable of continued expansion well after the host has entered stationary phase (Delbruck, 1946; Mollineux, 2006; Nguyen & Kang, 2014; Yin, 1993). These characteristics point to phage T7 being well-adapted to its strictly virulent lifestyle, in which phages infect and rapidly lyse the host, releasing dozens to hundreds of progeny per infection. This strategy is in stark contrast to the temperate phage lifestyle, in which the infecting phage has the option of integrating into the host chromosome (or, less commonly, maintaining itself as an episomal element) and passively replicating as part of the host genome until induced back into the lytic pathway at a later time (Gill & Hyman, 2010b). Part of T7's virulent replication strategy is driven by its single-subunit RNA polymerase, which recognizes unique T7 promoter elements and has extremely high processivity, features which have made it the backbone of modern protein expression strategies (Shis & Bennett, 2014). The presence of this conserved RNA polymerase is a hallmark for related T7-like podophages, and is a

defining of feature of the family *Autographiviridae*, of which T7 is the most famous member (Adriaenssens et al., 2020; Walker et al., 2020a).

Given T7's virulence, it is difficult to imagine how it could be "tamed" to pursue a temperate lifestyle and be safely maintained in the host genome. However, T7-like prophage elements have been observed within *Pseudomonas* and *Agrobacterium* genomes, but it has not been clear if these elements are intact and functional (Casjens, 2003; Dobbins et al., 2004; Mollineux, 2006). There have also been reports of cyanobacterial T7-like phage genomes encoding proteins associated with a temperate lifestyle (Sullivan, Coleman, Weigele, Rohwer, & Chisholm, 2005), but until very recently it was thought that these phages were unable to form functional lysogens. Recent reports suggest some of these T7-like cyanophages are capable of brief integration but not stable lysogen formation (Shitrit et al., 2021). More recently, there have been reports of T7-like phages capable of forming stable lysogens in *Pelagibacter* (Zhao et al., 2019) and T7-like prophage elements capable of induction and excision identified in *Pseudomonas putida* (Martínez-García, Jatsenko, Kivisaar, & de Lorenzo, 2015), providing evidence for the functionality of T7-like temperate phages and thus raising questions about the strictly virulent lifestyle classically associated with these phages. Truly temperate T7-like phages can provide ecological and evolutionary insight into how obligately lytic phages came to be and the origins of *Escherichia* phage T7, or more broadly its subfamily the *Studiervirinae* or family the *Autographiviridae*, which are characterized by podophage morphology and T7-like RNA polymerases. Here we identify two functional T7-like temperate phages infecting diverse Gram-negative hosts:

the phage ProddE, identified as a functional lysogen in the Deltaproteobacteria *Desulfovibrio desulfuricans*, and phage Pasto (Manuel et al., 2021), an environmentally-isolated phage infecting an *Agrobacterium* sp. in the Alphaproteobacteria. Using ProddE and Pasto as templates for temperate T7-like phage, we conduct a survey of extant bacterial genomes examining the widespread distribution of T7-like temperate phages within the domain *Bacteria*, suggesting that these are the ancestral form of *Studiervirinae* and perhaps many other members of the *Autographiviridae*.

5.2. Results and Discussion

ProddE and Pasto Are Functional Temperate Phages

ProddE (**Pro**phage of ***D***. ***d****esulfuricans* strain **E**delweiss) was initially identified following sequencing of another *Desulfovibrio* phage propagated on this host. Phage ProddE was assembled as an independent contig, indicating a significant level of spontaneous induction, replication and encapsidation in culture; analysis of the *D. desulfuricans* Edelweiss genome by PHASTER (Arndt et al., 2016) revealed the presence of two predicted incomplete prophages, one of which matched the assembled ProddE sequence. In the absence of a sensitive plating host for ProddE, alternative evidence for the functionality of this phage was examined. To assess whether excision of ProddE from the bacterial genome occurs, primers were designed which would only generate a product if phage *attL* and *attR* sites are brought together during prophage excision or during replication of concatemeric phage DNA (Martínez-García et al., 2015; Sozhamannan et al., 2006). The appropriately sized PCR product was obtained and

confirmed by Sanger sequencing to be the joining of the attL/attR site indicating that spontaneous phage induction and excision was occurring. Encapsidation of ProddE DNA was demonstrated by mapping sequence reads to assembled genomes, which showed that 96% (1,042,456) of mapped reads mapped to the primary *Desulfovibrio* phage being propagated in this culture**,** 4% (42,727) of reads mapped to ProddE, and <1% (855) of reads mapped to the host chromosome (**Figure 34**). This shows that ProddE was not simply assembled from contaminating host DNA but was enriched in the culture and protected from nuclease treatments associated with phage DNA purification. This evidence indicates ProddE is an intact prophage capable of excision, replication and virion production.

The *Agrobacterium* phage Pasto was isolated on its host *Agrobacterium* sp. strain AV1, referred to hereafter as *Agrobacterium* AV1, and sequenced as described previously (Manuel et al., 2021). Given that experimental evidence of T7-like temperate phage lysogeny is extremely limited (Martínez-García et al., 2015; Zhao et al., 2019) we sought to demonstrate true lysogeny (integration, excision, infection) in phage Pasto. Survivors from zones of Pasto clearing were isolated, and as expected of lysogens, lawns of these isolates displayed total immunity to Pasto infection with an efficiency of plating (EOP) of less than 10^{-9} relative to the parental strain **(Figure 35, left)**. Filter-sterilized culture supernatants from these isolates produced plaques when spotted on lawns of the parental *Agrobacterium* AV1 indicating spontaneous induction of phage Pasto in culture **(Figure 35, right)**. These results demonstrate Pasto is indeed a functional temperate

phage and add to the limited existing evidence that T7-like temperate phages are capable of both lytic and lysogenic life cycles.

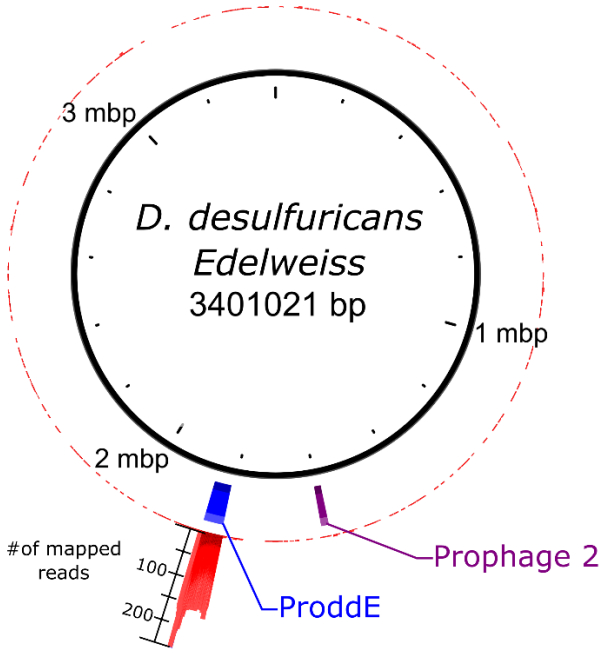


Figure 34. ProddE is an Intact Prophage.
Genomic map of *D. desulfuricans* Edelweiss with the locations of putative prophages ProddE and Prophage 2 as detected by PHASTER. Outer ring (red) displays Illumina short read mapping from the sequencing index in which ProddE was assembled, showing ProddE is preferentially encapsidated and is not simply contaminating host DNA reads, which map evenly at low coverage around the chromosome.

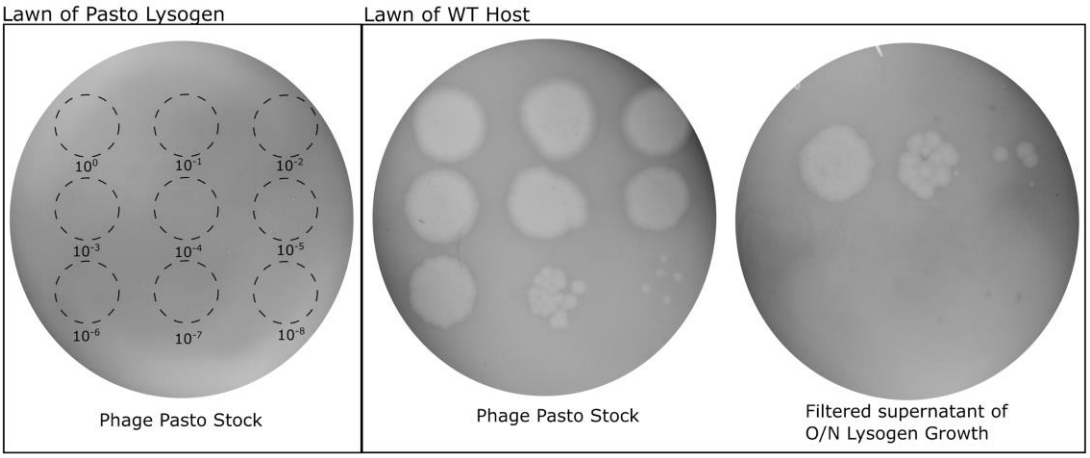


Figure 35. T7-like Temperate Phage Pasto Can Form Lysogens.
Plaque formation of Pasto on parental *Agrobacterium* AV1 (Center). Lysogens of *Agrobacterium* AV1 display complete immunity to phage Pasto (Left). Pasto lysogens display spontaneous prophage induction when the filtered overnight culture supernatant of a Pasto lysogen is plated on a lawn of the parental *Agrobacterium* AV1.

Comparison of ProddE and Pasto to Bacteriophage T7

ProddE and Pasto contain 47 and 60 protein-coding genes spanning their 42,637 bp and 42,407 bp genomes, respectively. These coding densities and genome sizes are comparable to that of T7 (NC_001604), which has 60 genes and a length of 39,937 bp. ProddE and Pasto contain 15 and 14 genes, respectively, with detectable similarity to those found in T7, including nearly all major replicative and structural proteins such as the RNA polymerase, helicase, DNA polymerase, portal, capsid, tail tubular proteins A and B, and the terminase large subunit, also called DNA maturase B (**Figure 36)**. Additionally, ProddE and Pasto contain a number of ORFs with no similarity to T7 proteins detectable by BlastP, but based on conserved domains, gene synteny, and size

likely serve similar functions as T7 proteins. Notably, the T7 internal virion proteins C (gp15) and D (gp16) have no recognizable similarity to proteins of the same size and location in ProddE or Pasto. These proteins are responsible for formation of the extended tail that is responsible for DNA translocation into the host (Hu, Margolin, Molineux, & Liu, 2013). These proteins physically interact with the host cell wall; therefore, this lack of amino acid sequence conservation may be explained by a necessity to maintain interactions with the cell walls of their respective hosts. Of note, outside the internal virion proteins, the only large (>350 aa) T7 gene products without detectable gene equivalents in both ProddE and Pasto are the T7 seryl-threonyl protein kinase (T7 gene *0.7*) and DNA ligase (T7 gene *1.3*).

The primary differences of the T7 and ProddE/Pasto gene repertoire are in small hypothetical proteins of unknown function that are distributed between major conserved proteins. Additionally, another set of differential proteins occupy an ~4.8 kb locus downstream of the large terminase in ProddE and Pasto containing the phage integrase and, in the case of Pasto, a putative lysogeny control gene. ProddE and Pasto share a large degree of genome synteny with T7, as can be seen in **Figure 36,** with both genomes being split predominantly into replicative and structural loci. Also similar to T7-like phages, all ProddE and Pasto genes are in a single orientation, with the exception of the predicted CI-like ProddE repressor directly upstream of the RNA polymerase. Repressors in temperate phages encoded in a divergent orientation from genes promoting the lytic cycle allowing for transcriptional control is common, as can be seen in phage lambda, *Lactobacillus* phage A2, and mycobacteriophages (Broussard et al., 2013;

Oppenheim, Kobiler, Stavans, Court, & Adhya, 2005). The structural gene regions of T7, ProddE, and Pasto have very few differences in terms of gene arrangement or presence of hypothetical proteins, compared to the DNA replication region which harbors a number of nonhomologous ORFs. Additional notable genes carried by ProddE and/or Pasto that are absent in T7 include putative lysogeny-related functions: both phages carry an integrase (tyrosine recombinase) and excisionase, while ProddE encodes a phage-like repressor and antirepressor, and Pasto encodes a MarR-like transcriptional regulator. Additionally, ProddE and Pasto also encode an ArdA-like anti-restriction protein (Belogurov, Delver, & Rodzevich, 1993). ArdA likely carries out a similar function to gene *0.3* of T7 which encodes the Ocr protein (Dunn, Elzinga, Mark, & Studier, 1981), and although ArdA has been described to be similarly acidic and serve a similar role as the Ocr protein (Belogurov et al., 1993), the ProddE and Pasto ArdA-like proteins contain no amino acid sequence similarity to Ocr of T7. ProddE and Pasto are displayed in their integrated prophage forms in **Figure 36**, in order to retain synteny with other identified T7-like temperate phages discussed below. In an infectious virion, the first early genes into the host are predicted to be those downstream of DNA maturase B (**Figure 36, right side**). In the case of ProddE, this early gene region would include the excisionase, ArdA-like protein, integrase, and the repressor. Conservation of gene content and synteny suggests a distant but common origin of ProddE, Pasto, and T7.

In addition to gene synteny, ProddE and Pasto share some terminator locations with T7 as indicated in **Figure 36**. T7 has two terminators that facilitate its unique multistep DNA ejection process (Mollineux, 2006). Both ProddE and T7 have

transcriptional terminators in close proximity to RNA polymerase and immediately following the major capsid protein. In contrast, Pasto has a putative terminator upstream of RNA polymerase, but still shares the conserved terminator with T7 and ProddE immediately distal to the gene for the major capsid protein. The conservation of these terminators suggests ProddE and Pasto follows a similar transcriptional program to that of T7. In addition to these two conserved terminators, ProddE has an additional five predicted terminators, some of which fall within the predicted ProddE lysogeny cassette, suggesting roles in the lysis/lysogeny decision-making process, while Pasto contains an additional two predicted transcriptional terminators. ProddE and Pasto's genome synteny and relatedness to T7 suggests that transition from the temperate to the virulent state does not necessarily require fundamental genomic changes, and these phages may be able to move between these lifestyles by addition or removal of relatively small, discreet gene modules; this also raises the possibility that the highly virulent phage T7 and its relatives may have arisen from a temperate ancestor.

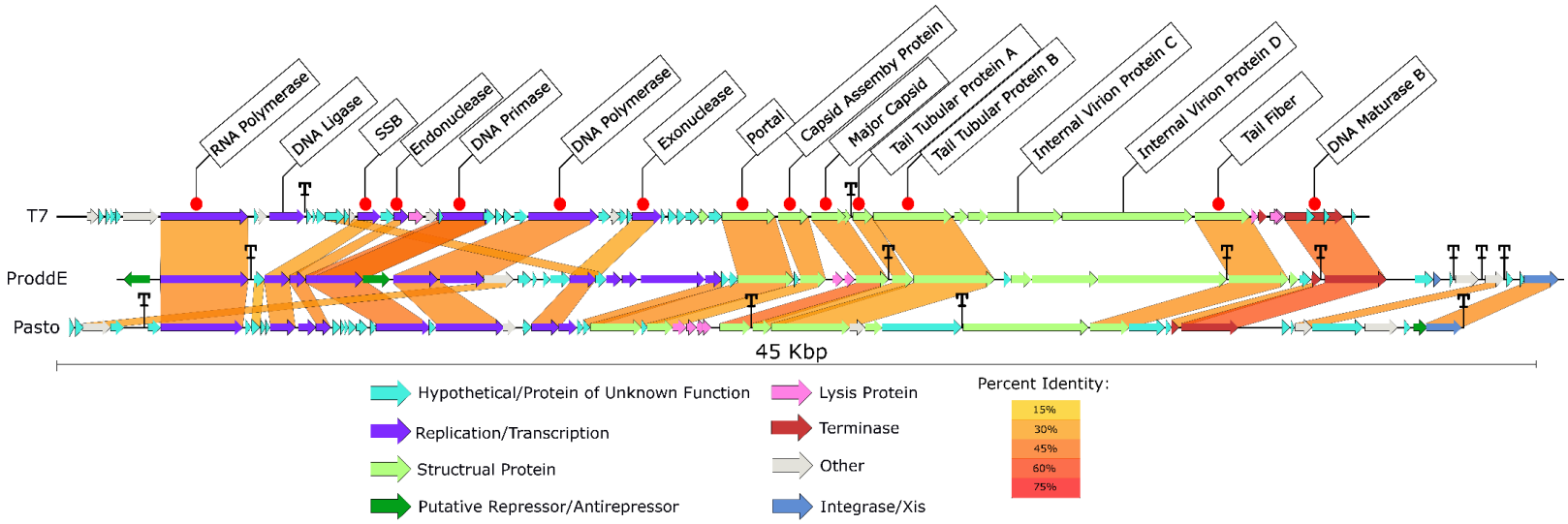


Figure 36. Similarity of ProddE and Pasto to Phage T7.
Genomic comparisons of T7, ProddE, Pasto. Gene homologs with an evalue <1e-5 are connected by bars to their respective neighbor with color scale indicating percent identity as determined by BLASTp. To display conserved synteny phage genomes are aligned to their RNA polymerase gene. Pasto was reopened after the integrase to be syntetic with ProddE and other prophage elements below. Gene classifications are colored as indicated in the legend. Red dots above T7 genes indicate genes conserved in all representative genomes. “T”’s annotated on genomes indicate canonical T7 terminators (NC_001604.1), and putative ProddE and Pasto terminators as predicted by TransTermHP.

Temperate T7-like phages are widespread across the domain Bacteria and suggest a common ancestor

We wished to determine how common ProddE, Pasto, and other reported T7-like temperate phages (Casjens, 2003; Mollineux, 2006; M. B. Sullivan et al., 2005; Zhao et al., 2019) are in other bacterial genomes. BLASTn searches of bacterial genomes based on nucleotide similarity to ProddE, Pasto or T7 returned few matches that only aligned to small portions of the query sequences. To conduct searches based on protein sequence, we used tBLASTn to search for discrete chromosomal regions containing homologs of the T7 RNA polymerase, DNA polymerase, tail tube B, and large terminase (maturase B) as potential T7-like prophage elements in bacterial genomes. This search yielded 45 unique elements that met the criteria of containing these four T7 homologs within a ~50 kb region **(Table 19, Table 20)**. Seemingly intact T7-like temperate phages with a high degree of gene synteny to ProddE, Pasto and T7 **(Figure 37)** were found integrated into organisms from across diverse bacterial genera, including Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, and, with the identification of ProddE, Deltaproteobacteria. The only class within the Proteobacteria lacking an identified T7-like temperate phage was Epsilonproteobacteria. Outside of the Proteobacteria, two T7-like elements were identified in the class Negativcutes in the phylum Firmicutes. It should be noted a number of T7-like temperate prophage elements in the PhiKMV genomic arrangement were also identified, however they failed to meet the strict criteria for a T7-like phage utilized in this study and were thus excluded from

further analysis (**Figure 38**). While these T7-like phages are widely distributed across bacterial phyla, notably no T7-like prophage elements meeting our criteria were identified within the Enterobacteriaceae despite this family being highly represented within deposited bacterial genomes.

To investigate the relatedness of these prophage elements spanning different classes, a representative T7-like prophage element was selected from each class based on its nucleotide identity to T7 and compared to its neighbor by BLASTp (**Figure 39**). As was the case with ProddE, Pasto, and T7, each genome is divided into replicative and structural loci, all genes are predominantly in a single orientation, and are syntenic to the T7 gene arrangement. The "core" genome region, spanning the RNA polymerase to maturase B, is largely conserved in order and content, with flanking regions encoding multiple hypothetical proteins and, in the case of the temperate phages or predicted prophage elements, integrases and other lysogeny functions. As was the case with T7, ProddE, and Pasto, the same 13 genes are conserved between T7 and all the other T7-like temperate phages compared in **Figure 39,** denoted by red dots. As shown in **Figure 40**, the RNA polymerase, DNA primase, DNA polymerase, exonuclease, portal, capsid, tail tube AB, maturase B and tyrosine recombinase integrase are nearly universally conserved in ProddE, Pasto and the 45 T7-like predicted prophage elements. The ArdA-like anti-restriction protein and endonuclease are detectable in ~45% of genomes. These conserved genes could constitute the ancestral "core" genes from which T7-like phages and other *Autographiviridae* have originated. Like ProddE and Pasto, there was low conservation in internal virion proteins and phage tail fiber proteins, again likely to

maintain interactions with diverse hosts. Also, notably absent from the majority of other T7-like prophage elements are putative repressors and antirepressors homologous to those identified in ProddE (Gene 1 and Gene 7 in **Figure 40**), suggesting these elements may employ diverse mechanisms to establish and maintain lysogeny.

In addition to conserved synteny and gene content, many of the T7-like prophage elements have conserved insertion sites in their respective bacterial genomes. Of the 45 T7-like predicted prophage elements surveyed, 16 (36%) were inserted downstream of genes encoding host Clp protease and trigger factor with or without a recognizable $tRNA_{leu}$, consistent with the insertion site observed in temperate T7-like *Pelagibacter* phages (Zhao et al., 2019). This conservation of insertion region is not restricted to prophage elements of the same genus, but span hosts from the Alpha-, Beta-, Gamma-, and Deltaproteobacteria.

Host Class	T7-like Prophages Identified	Unique T7-like prophage elements	Approximate sequences queried in NCBI nt	Identified prophages / Sequences queried
Alphaproteobacteria	11	10	320k	.003%
Betaproteobacteria	10	9	175k	.006%
Gammaproteobacteria	41	22	690k	.006%
Deltaproteobacteria	1(ProddE)	1	60k	.002%
Negativicutes	2	2	10k	.02%

Table 19. Summary of T7-like Prophages Identified in Bacterial Chromosomes. Number of unique T7-like prophages by host class identified in the Domain Bacteria (taxId:2), including ProddE. Only (43/45) and (64/66) are shown in this table; two regions were identified in taxonomically unclassified organisms. In the right column, the number of prophages identified is normalized to the number of sequences searched, as

reported by BLAST at runtime. No T7-like elements were identified in the Epsilonproteobacteria.

Host Class	Accession	Organism	DNA Polymerase Start	Large Terminase Start	RNA Polymerase Start	Tail Tubular B Start
Unknown	MH327397.1	MH327397.1 Bacterium AG-345-A23 Ga0172254_11 genomic sequence	20247	6428	15306	33526
Alphaproteobacteria	CP021219.1	CP021219.1 Sinorhizobium meliloti RU11/001 chromosome, complete genome	3418186	3395952	3426525	3408490
Alphaproteobacteria	CP014260.1	CP014260.1 Agrobacterium tumefaciens strain S33 chromosome linear, complete sequence	2954030	2977284	2946638	2964665
Gammaproteobacteria	CP041967.1	CP041967.1 Xanthomonas citri pv. glycines strain K2 chromosome, complete genome	1296421	1273210	1303988	1286459
Alphaproteobacteria	LT671861.1	LT671861.1 Ochrobactrum anthropi genome assembly, chromosome: I	1091617	1067738	1098899	1081309
Alphaproteobacteria	CP026525.1	CP026525.1 Sinorhizobium meliloti strain AK21 chromosome, complete genome	1731607	1754332	1723013	1741506
Gammaproteobacteria	CP014014.1	CP014014.1 Stenotrophomonas maltophilia strain FDAARGOS_92 chromosome, complete genome	4808834	4784030	4815469	4798434
Betaproteobacteria	CP027773.1	CP027773.1 Variovorax sp. PMC12 chromosome 1, complete sequence	1540982	1563672	1531938	1550322
Alphaproteobacteria	CP051630.1	CP051630.1 Pyruvatibacter mobilis strain CGMCC 1.15125 chromosome, complete genome	87129	109945	78028	96583
Betaproteobacteria	CP064338.1	CP064338.1 Schlegelella thermodepolymerans strain DSM 15344 chromosome, complete genome	2559853	2584959	2553033	2569400
Gammaproteobacteria	CP011010.1	CP011010.1 Stenotrophomonas maltophilia strain ISMMS3, complete genome	1006571	982181	1014325	996434
Gammaproteobacteria	CP027710.1	CP027710.1 Pseudomonas chlororaphis subsp. piscium strain SLPH10 chromosome, complete genome	2049049	2025152	2057344	2037299
Gammaproteobacteria	CP048878.1	CP048878.1 Spartinivicinus ruber strain S2-4-1H chromosome, complete genome	1144149	1120068	1152591	1133210
Gammaproteobacteria	CP049956.1	CP049956.1 Stenotrophomonas maltophilia strain NCTC10498 chromosome, complete genome	379410	354624	387164	369016
Gammaproteobacteria	CP044092.1	CP044092.1 Stenotrophomonas maltophilia strain FDAARGOS_649 chromosome, complete genome	2263420	2238622	2271174	2253026
Gammaproteobacteria	CP040440.1	CP040440.1 Stenotrophomonas maltophilia strain ICU331 chromosome, complete genome	1004328	979528	1012082	993935
Gammaproteobacteria	AP021908.1	AP021908.1 Stenotrophomonas maltophilia DNA, complete genome, strain: WP1-W18-CRE-01	1007742	982938	1015496	997342
Gammaproteobacteria	CP040436.1	CP040436.1 Stenotrophomonas maltophilia strain PEG-390 chromosome, complete genome	963293	938487	971053	952900
Negativicutes	LR215982.1	LR215982.1 Phascolarctobacterium faecium isolate Phascolarctobacterium succinatutens 82G5 genome assembly, chromosome: 1	1963195	1939761	1972394	1952204
Alphaproteobacteria	HG938353.1	HG938353.1 Neorhizobium galegae, complete genome	1319849	1342651	1309927	1329232
Gammaproteobacteria	AE015451.2	AE015451.2 Pseudomonas putida KT2440 complete genome	2595165	2619497	2586675	2605491
Gammaproteobacteria	CP015876.1	CP015876.1 Pseudomonas putida SJTE-1, complete genome	3978715	3954086	3987242	3968399
Betaproteobacteria	CP035913.1	CP035913.1 Massilia lutea strain DSM 17473 chromosome	7367808	7342230	7375437	7356599
Alphaproteobacteria	AP009384.1	AP009384.1 Azorhizobium caulinodans ORS 571 DNA, complete genome	4129436	4104692	4137966	4117054
Unknown	MH327293.1	MH327293.1 Bacterium AG-316-A11 Ga0172227_11 genomic sequence	35670	22220	30219	2098
Betaproteobacteria	CP013448.1	CP013448.1 Burkholderia sp. Bp7605 strain MSMB0175 chromosome 1, complete sequence	1092468	1119017	1085426	1101714
Betaproteobacteria	CP008727.1	CP008727.1 Burkholderia oklahomensis strain EO147 chromosome 2, complete sequence	733843	760380	726599	743171
Betaproteobacteria	CP033704.1	CP033704.1 Burkholderia pseudomallei strain FDAARGOS_593 chromosome 2, complete sequence	2223564	2196730	2230744	2213688
Gammaproteobacteria	CP014037.1	CP014037.1 Vibrio diabolicus strain FDAARGOS_105 chromosome 2, complete sequence	1198041	1170193	1204434	1183009
Gammaproteobacteria	CP051110.1	CP051110.1 Vibrio alginolyticus strain 2015AW-0011 chromosome 2	1094890	1067039	1101283	1079855
Gammaproteobacteria	CP017890.1	CP017890.1 Vibrio alginolyticus strain K01M1 chromosome 2, complete sequence	832602	860640	826210	847638
Gammaproteobacteria	CP007151.1	CP007151.1 Marinobacter similis strain A3d10, complete genome	3659128	3682188	3647034	3668329
Betaproteobacteria	CP014844.1	CP014844.1 Cupriavidus nantongensis strain X1 chromosome 1, complete sequence	3849541	3876220	3840893	3858987
Betaproteobacteria	CP013418.1	CP013418.1 Burkholderia sp. MSMB0266 chromosome 2, complete sequence	1658826	1631728	1667060	1649382
Gammaproteobacteria	CP045324.1	CP045324.1 Marinobacter sp. THAF197a chromosome, complete genome	2608062	2585069	2620551	2599129
Alphaproteobacteria	CP050898.1	CP050898.1 Rhizobium pusense strain FDAARGOS_633 chromosome 1	573670	547880	583501	565282
Gammaproteobacteria	CP025999.1	CP025999.1 Stenotrophomonas sp. SAU14A_NAIMI4_8 chromosome	1034189	1061221	1025534	1046266
Alphaproteobacteria	LR743509.1	LR743509.1 Hyphomicrobium sp. ghe19 isolate hypp genome assembly, chromosome: 1	2602731	2627192	2591319	2612707
Betaproteobacteria	LT977070.1	LT977070.1 Cupriavidus taiwanensis isolate Cupriavidus taiwanensis SWF 65033 genome assembly, chromosome: CBM2638_a	2784207	2811160	2775235	2793940
Gammaproteobacteria	CP002470.1	CP002470.1 Vibrio vulnificus MO6-24/O chromosome II, complete sequence	667510	697187	661191	684092
Gammaproteobacteria	AP019656.1	AP019656.1 Vibrio panuliri JCM 19500 plasmid pVPan1 DNA, complete sequence	33970	7275	27713	43580
Alphaproteobacteria	CP021030.1	CP021030.1 Rhizobium sp. NXC14 chromosome, complete genome	1972166	1999083	1962736	1987370
Gammaproteobacteria	CP028899.1	CP028899.1 Stenotrophomonas maltophilia strain AB550 chromosome, complete genome	997118	1024822	988188	1008576
Negativicutes	CP011940.1	CP011940.1 Megasphaera hexanoica strain MH chromosome, complete genome	1488719	1516828	1478810	1499730
Gammaproteobacteria	CP026055.1	CP026055.1 Aeromonas caviae strain FDAARGOS_72 chromosome, complete genome	1281631	1250749	1289609	1266329

Table 20. Unique Prophages Elements Identified in NR/NT Database

The 45 unique prophage elements identified in the nr/nt database as described in the materials/methods. Indicated host class to which the host belongs, accession number, ID title, nt coordinate within the bacterial chromosome in which DNA polymerase, RNA polymerase, Large Terminase, Tail Tubular start coordinates.

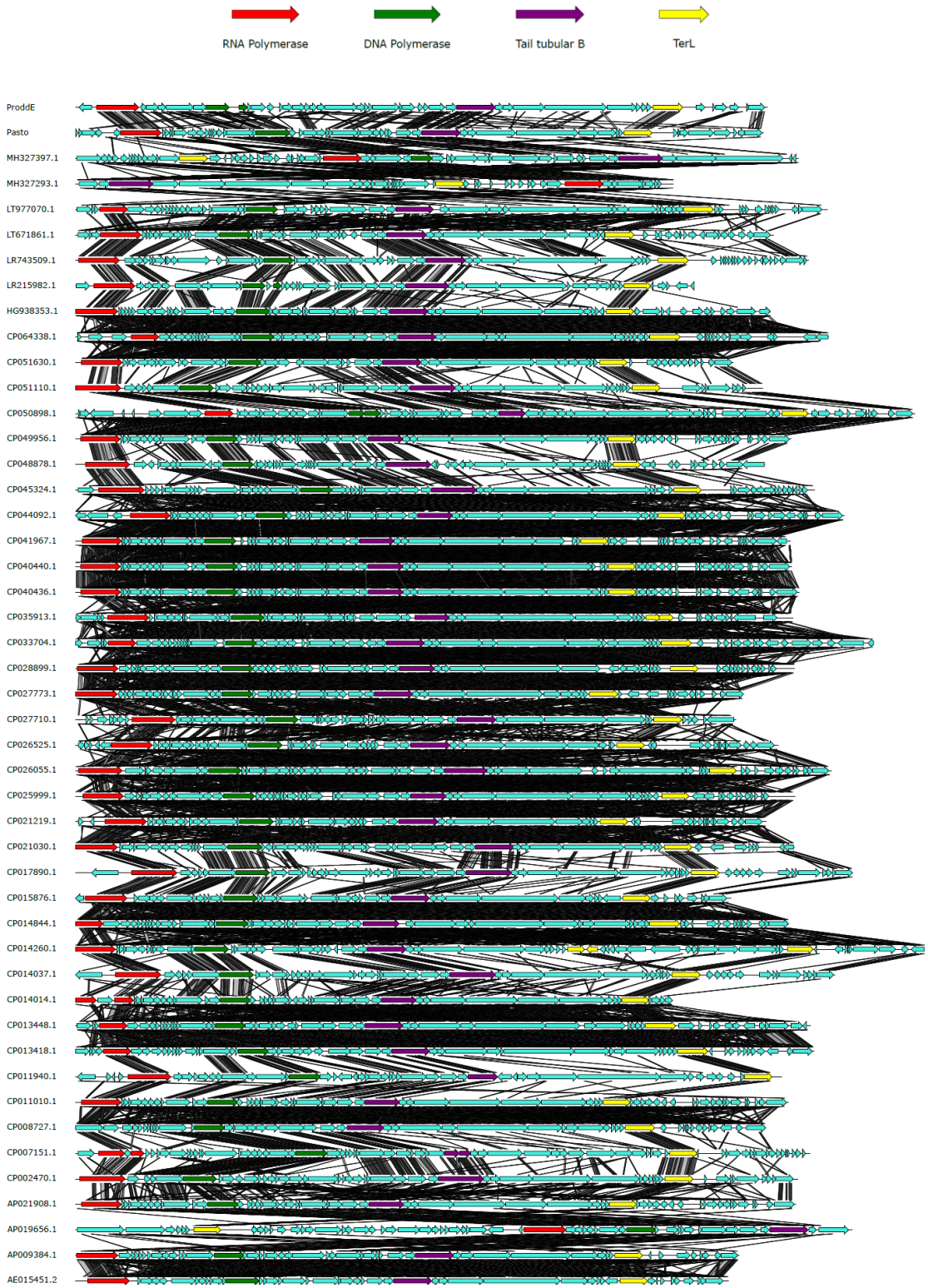


Figure 37. Genome Synteny of T7-like Temperate Elements

EasyFig (tblastx) of the Prokka automatic annotation of 45 identified T7-like temperate elements identified in the nr/nt database, as well as ProddE and Pasto. Conserved genome synteny of major proteins can be seen, the T7-protein homolog used to identify these phages from Genbank are colored accordingly to assist in visualization. The conserved genome arrangement can be seen in the order (RNA polymerase(red)-DNA polymerase(green)-Tail Tubular Protein B(purple)-DNA Maturase B(yellow). In the few seeming exceptions, M327397.1, MH327293.1, and AP019656.1 genomic synteny is still maintained, however these elements are deposited as their own contig or listed as plasmids and simply opened at a different coordinate. MH327293.1 contains a DNApol unannotated by Prokka likely due to its proximity to the end of phage contig, however tblastx lines drawn by EasyFig indicating homology from other DNApols can be seen.

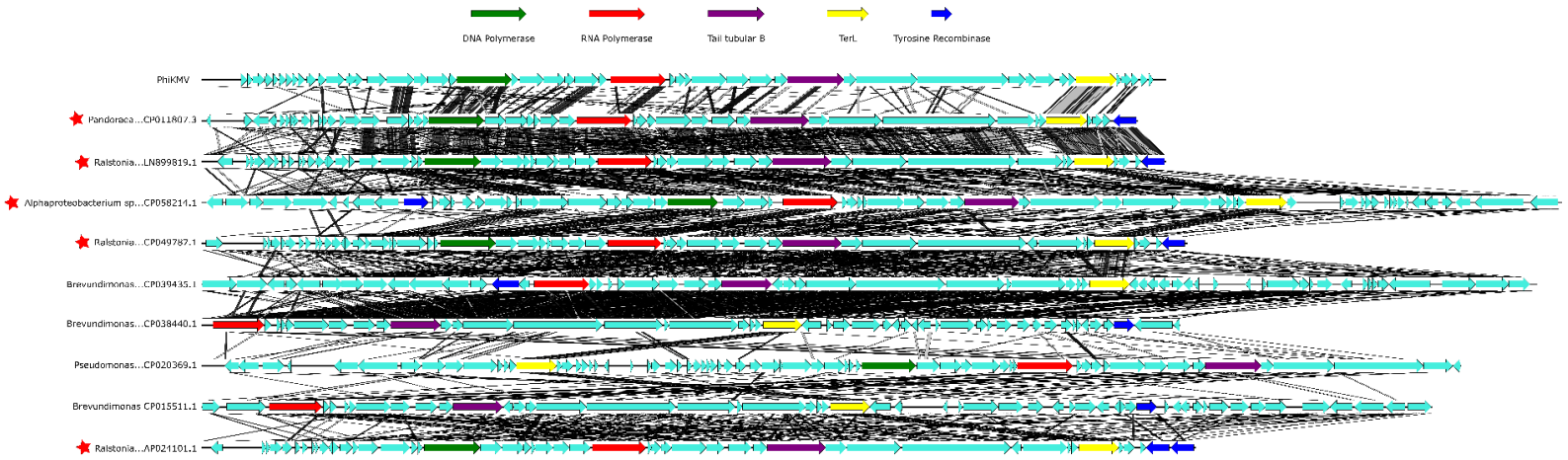


Figure 38. Putative PhiKMV-like Prophage Elements Also Identified.
EasyFig (tblastx) of the Prokka automatic annotation of additional prophage elements identified that did not meet criteria. These elements seemingly represent PhiKMV orientations of T7-like temperate prophage genome. PhiKMV included as reference (top). The conserved genome arrangement can be seen in the order (RNA polymerase(red)-DNA polymerase(green)-Tail Tubular Protein B(purple)-DNA Maturase B(yellow)-Tyrosine Recombinase (blue) particularly those denoted with a red star, which most closely resemble PhiKMV.

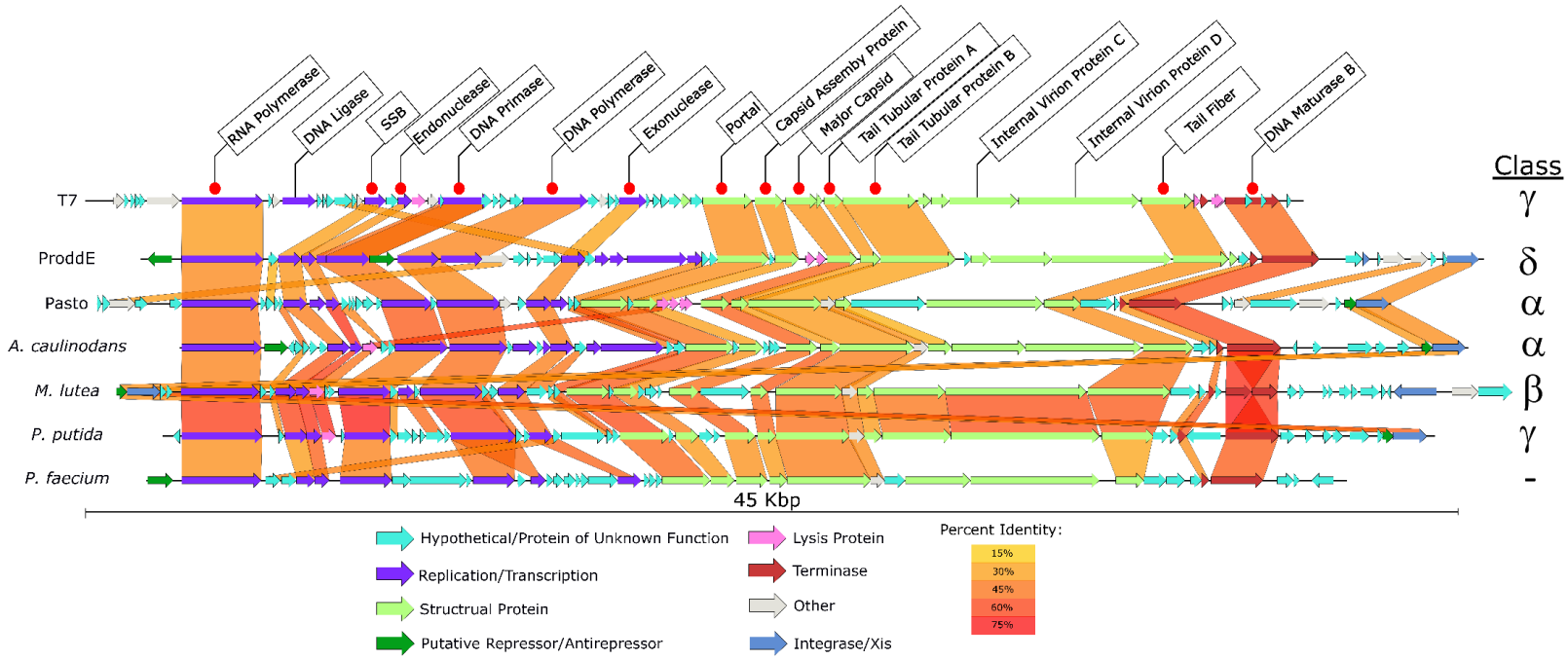


Figure 39. Conservation of ProddE, Pasto, and other T7-like Prophage Elements.
Genomic comparisons of T7, ProddE, Pasto as in Figure 3, with addition of representative T7-like prophage elements identified from diverse hosts (AP009384.1, CP035913.1, LR215982.1, and CP015876.1). To display conserved synteny phage genomes are aligned to their RNA polymerase gene, and prophage elements are displayed in their integrated forms. Gene classifications are colored as indicated in the legend. Left, genus and species in which respective prophage element was identified. Labels on the right side of the figure indicate the class of each phage's respective host (Gamma-, Delta-, Alpha-, Beta- Proteobacteria, and Negativicutes). Red dots above T7 genes indicate genes conserved in all representative genomes.

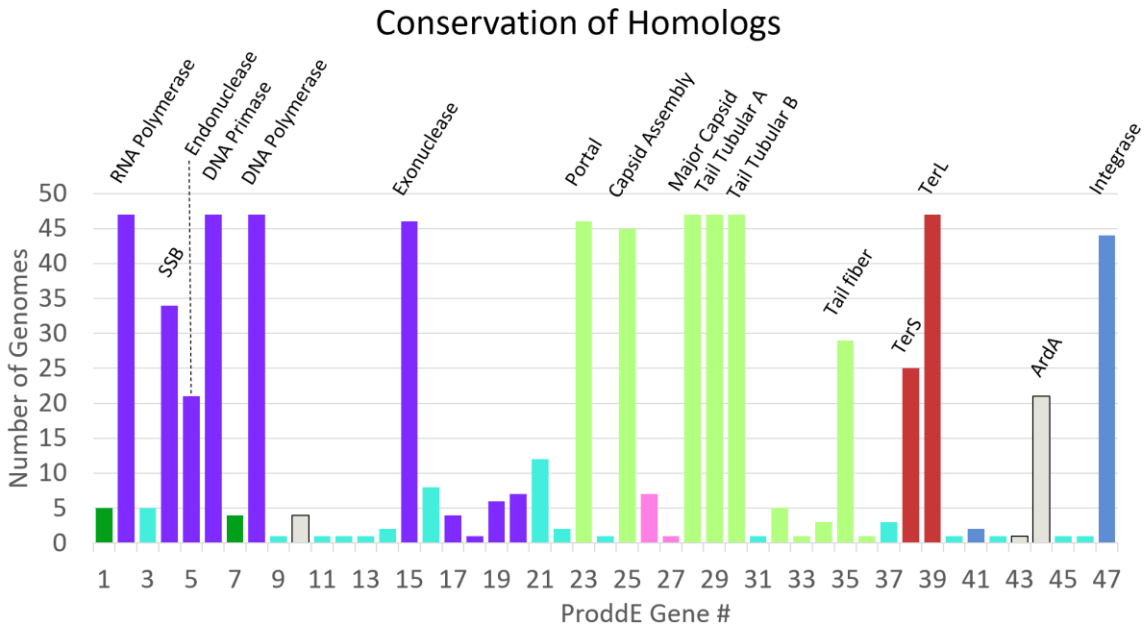


Figure 40. Conservation of Gene Homologs Amongst Identified T7-like Temperate Phages.
Using ProddE proteins as query, a tblastn search was conducted against all identified prophage elements, Pasto, and ProddE (47 genomes total). Genes conserved in >40% of genomes are shown with a label. Bar color codes correlate to figure legend in Figures 3 and 4.

Relationships of T7-like prophage elements to the broader *Autographiviridae*

Variable mutation rates, genome mosaicism, and differing life styles complicate phylogenetic calculations in phages (Adriaenssens et al., 2020; Casjens & Thuman-Commike, 2011; Lawrence, Hatfull, & Hendrix, 2002). In the dataset of T7-like temperate phages described here, nucleotide identity is nearly absent between phages from different host classes and protein sequence identity is low. Most of the conserved proteins of the T7-like temperate phages that would be ideal for multiple sequence

alignment-based phylogenetic analysis like the RNA polymerase, DNA polymerase, tail tubular protein B, and TerL share only an average of 36%, 43%, 33%, and 54% protein sequence identity, respectively, and considerably lower identity among prophage elements occupying hosts of different classes. These low identities approach the "twilight zone" in which methods that rely on multiple sequence alignments become unreliable (Bhardwaj et al., 2012). In order to further study the diversity and relationships of these T7-like temperate elements, vContact2 was used to compare phages based on total protein-coding gene content (Ho Bin Jang et al., 2019; Bolduc et al., 2017). vContact2 was originally designed for classifying metagenomic sequences into taxonomic clusters based on gene sharing networks through formation of protein clusters with a reference database. The resulting output places genomes into viral clusters and sub-clusters, which can subsequently be visualized as phage networks in which node proximity equates to phage similarity based on protein content. For this study we applied vContact2 with the NCBI Bacterial and Archaeal Viral RefSeq v88 database to the 45 identified unique T7-like prophage elements, ProddE, and Pasto, with the known virulent *Autographivirdae* T7, PhiKMV, SP6, and P60 included as internal controls. Visualization of resulting network output is shown in **Figure 41.** Viral clusters formed by vContact2 are generally in agreement with the ICTV taxonomic classifications (Walker et al., 2020a). As expected, all T7-like temperate phages fell within the broader *Autographiviridae* group (**Figure 41A, inset**). With the exception of a few outliers (teal triangles) that did not meet the vContact2 cutoff score, all the T7-like temperate phages were placed within a single cluster (maroon triangles). Also within this

cluster are 22 phages from the RefSeq database that had been previously isolated as virions (maroon squares). These include a number of cyanobacterial phages infecting *Prochlorococcus* and *Synechococcus* genera, as well as *Mesorhizobium* and *Pelagibacter* phages **(Figure 41A)(Table 21).** Inspection of these phages showed them to be T7-like phages with varying content of proteins required for lysogeny. As expected, these also include the previously mentioned cyanophages (Sullivan et al., 2005) and *Pelagibacter* phages (Zhao et al., 2019) which contain integrases. However, 12 of these clustered phages from RefSeq lack integrases. It is possible these phages lacking integrases and/or other lysogeny proteins represent intermediates between the temperate and virulent lifestyle. Since the input T7-like temperate phage elements clustered with additional diverse *Autographiviradae* members within RefSeq, we designated this cluster as the Prophage elements + *Autographiviridae* cluster, abbreviated **PeA**. Based on ICTV taxonomy (Adriaenssens et al., 2020; Walker et al., 2020a), many of the RefSeq phages within the PeA cluster are unclassified, or constitute their own genus/species of which they are the sole member.

Given the clustering of cyanophages with the T7-like temperate elements, it is striking that none of the 45 identified T7-like prophage elements were found in any cyanobacterial genomes. However, our search was restricted to a conservative definition of what makes a phage "T7-like" (similar-sized genome, and RNApol, DNApol, tailtube B, and DNA maturase B with detectable similarity to T7 homologs). It is also possible T7-like cyanophages only form transient lysogens, as has recently been suggested (Shitrit et al., 2021), making capturing of integrated prophages in genomic sequences

challenging. To investigate this further, the T7 RNA polymerase was specifically queried against the NCBI nr/nt database restricted to Cyanobacteria (taxid:1117) via tBLASTn. This yielded six hits, which showed T7-like RNA polymerases in proximity to an integrase, but beyond this, no recognizable elements that would indicate the presence of intact prophages. This may indicate stably integrated T7-like prophages in these Cyanobacteria genomes have since degraded, but confirms the above analysis that identifies these elements as residing primarily in proteobacterial hosts.

The diversity of host species of the phage in the PeA cluster is notably much greater than the other major clusters of *Autographiviridae*, as can be seen in **Figure 41B,** spanning host classes from the Alpha-, Beta-, Gamma-, and Delta-proteobacteria, as well as the Negativicutes and Cyanophyceae**.** The rest of the *Autographiviridae* outside of the PeA cluster are overwhelmingly restricted to the Gammaproteobacteria **(Figure 41B)**. Despite being a distinct cluster as indicated by vContact, it is clear that the PeA nodes are much more diffuse, and even overlap with virulent nodes in some cases, such as the *Studiervirinae* (red) or *Molineuxvirinae* (blue). This broader clustering of PeA can likely be explained by the diversity of hosts to which these temperate phages have become adapted (Bobay, Rocha, & Touchon, 2013). The formation of this diffuse PeA cluster distinct from other prevalent subfamilies of the *Autographiviridae* suggests a very distant common origin.

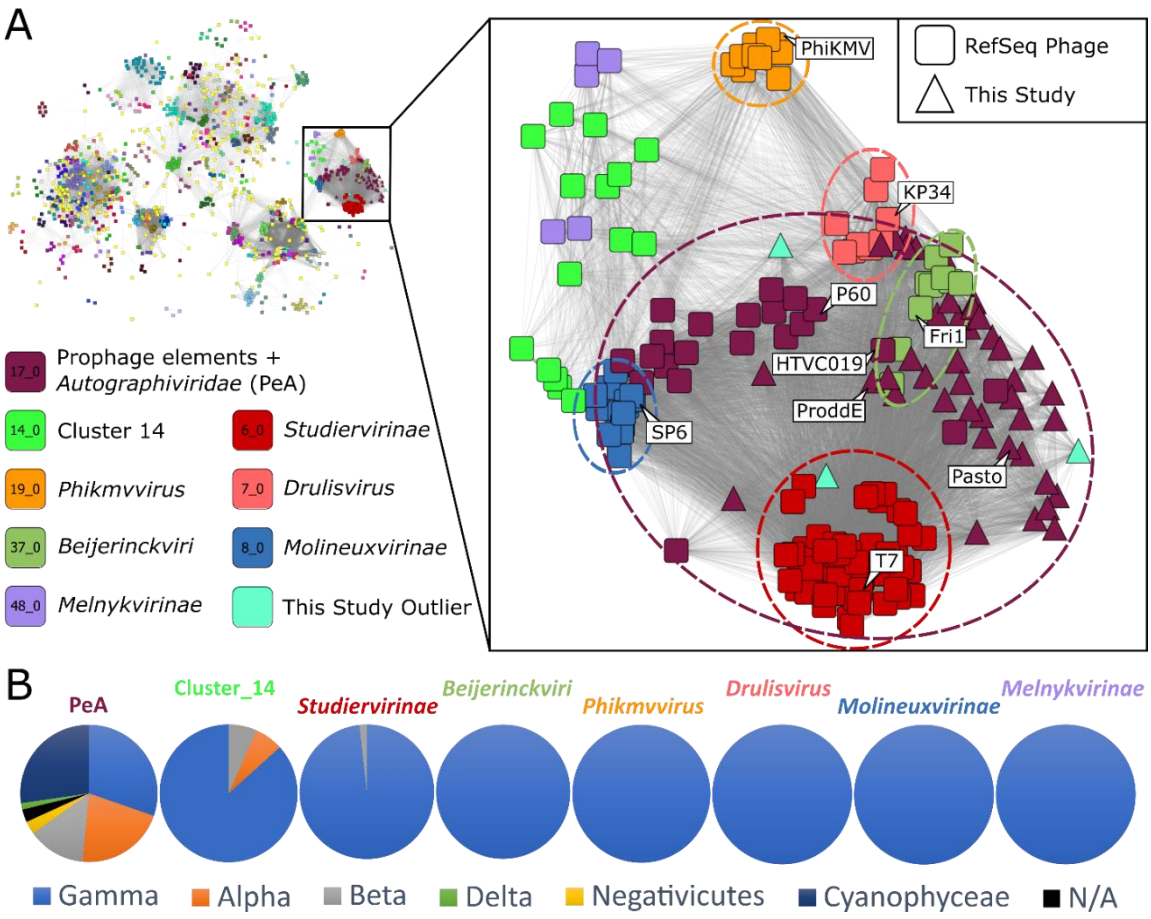


Figure 41. Taxonomic Placement and Clustering of T7-like Temperate Elements Amongst *Autographiviridae*.
A) vContact2 was used to generate phage genome clusters from the RefSeq database plus 45 identified T7-like prophage elements, ProddE, and Pasto; results were visualized in Cytoscape. Increasing edge opacity correlates to increasing phage similarity. In the upper left, large portion of the clustering network map containing the *Caudovirales* order is shown, with the cluster containing *Autographiviridae* and novel T7-like temperate phages enlarged. Node shapes and colors are indicated in the figure legend. Cluster numbers as assigned by vContact are indicated within the legend node. For context, clusters were assigned a name based on highest ICTV taxonomic classification which

encompasses phages that dominate the cluster, with the exception of cluster 14, which was too heterogeneous to assign an accurate taxonomic grouping. Founding members of their respective clusters, or specific phage nodes discussed in this work are noted as callouts. To assist visualization of clusters, boundaries are drawn around relevant clusters and RefSeq unclustered outliers are removed from the visualization. B.) Diversity of phage hosts are shown by cluster. Most of the RefSeq *Autographiviridae* phages in the major genera utilize gammaproteobacterial hosts. In contrast, the phages in the PeA cluster, which contains predicted prophage elements, ProddE and Pasto (maroon triangles) and other unclassified *Autographiviridae* (maroon squares) utilize a much more diverse set of hosts.

Genome	Family	VC
Cyanophage~9515-10a	Podoviridae	17_0
Cyanophage~KBS-P-1A	Podoviridae	17_0
Cyanophage~NATL1A-7	Podoviridae	17_0
Cyanophage~NATL2A-133	Podoviridae	17_0
Cyanophage~P-SSP2	Podoviridae	17_0
Cyanophage~SS120-1	Podoviridae	17_0
Mesorhizobium~phage~vB_MloP_Lo5R7ANS	Podoviridae	17_0
Pelagibacter~phage~HTVC011P	Podoviridae	17_0
Pelagibacter~phage~HTVC019P	Podoviridae	17_0
Podovirus~Lau218	Podoviridae	17_0
Prochlorococcus~phage~P-GSP1	Podoviridae	17_0
Prochlorococcus~phage~P-SSP10	Podoviridae	17_0
Prochlorococcus~phage~P-SSP3	Podoviridae	17_0
Prochlorococcus~virus~PSSP7	Podoviridae	17_0
Synechococcus~phage~S-CBP1	Podoviridae	17_0
Synechococcus~phage~S-CBP2	Podoviridae	17_0
Synechococcus~phage~S-CBP3	Podoviridae	17_0
Synechococcus~phage~S-CBP4	Podoviridae	17_0
Synechococcus~phage~S-RIP1	Podoviridae	17_0
Synechococcus~phage~S-RIP2	Podoviridae	17_0
Synechococcus~virus~P60	Podoviridae	17_0
Synechococcus~virus~Syn5	Podoviridae	17_0

Table 21. Additional Viral Genomes Clustered with T7-like Temperate Elements. Subsection of genome by genome overview from vContact output. Table describes viral genomes from NCBI Bacterial and Archaeal Viral RefSeq v88 database which co-cluster with T7-like temperate elements forming the PeA cluster.

Maintenance of lysogeny and integration

Well-studied temperate phages such as coliphage lambda regulate the lysis-lysogeny decision and lysogen maintenance via a genetic "switch" (Oppenheim et al., 2005). These switches typically consist of divergent transcriptional units controlled by overlapping promoter-operator elements, and transcriptional activators and repressors

encoded by these transcripts (Brüssow & Desiere, 2001). The general absence of conserved phage-like repressors in the T7-like prophage elements prompted a deeper exploration of how these elements might be maintaining their lysogenic states. Phage Pasto, which also lacks an obvious phage-like repressor, has been shown to form stable lysogens (**Figure 35**), thus these phages might use a novel strategy to establish and control the lysogenic state. To explore this, vContact2 protein clusters were investigated for putative lysogeny control proteins conserved in the T7-like temperate phages and prophage elements. Only 5 of the 47 identified T7-like temperate elements harbored a CI-like repressor that fell within the same protein cluster as the repressor of ProddE with variable presence of a similar ProddE antirepressor. Identified CI-like repressors in the T7-like temperate phages are restricted to the Beta- and Deltaproteobacteria. It is also worth noting that none of the other RefSeq phage genomes within the PeA cluster contained homologs to the repressor or antirepressor identified in ProddE.

The predicted lysogenic switch in ProddE is a pair of divergent transcripts that encode a phage CI-like repressor with a LexA-like cleavage domain (cd06529) on the minus strand and the phage RNA polymerase on the plus strand, driven by predicted overlapping promoters (**Figure 42A**). This suggests a mechanism of prophage induction that is sensitive to the bacterial SOS response as is observed in Lambda (Galkin et al., 2009). The position of a recognizable phage-like repressor upstream of the phage RNA polymerase provides a plausible mechanism for the lysogenic repression of a T7-like phage genome, which is otherwise considered to be highly virulent. All mid- and late gene expression in phage T7 is dependent on the expression of the phage RNA

polymerase, which recognizes phage-specific promoters that are otherwise not recognized by the host RNA polymerase (Mollineux, 2006). Because of this dependence, repression of the phage RNA polymerase would prevent expression of the phage middle (Class II) and late (Class III) transcripts, which include viral DNA replication, viral morphogenesis, host DNA degradation, and lysis functions (Mollineux, 2006), and allow the phage to exist as a quiescent prophage element within the cell. While phage ProddE encodes a recognizable phage-like repressor upstream of the phage RNA polymerase gene, the more common putative lysogeny control mechanism identified in the T7-like temperate phage elements is a group of MarR-like transcriptional regulators. These are present in 10 of the 47 T7-like temperate elements, including *Agrobacterium* phage Pasto, and additionally in 9 of the RefSeq phage genomes in the PeA cluster shown in **Figure 41**. Within the PeA cluster, MarR-like repressor is present in the Cyanobacteria as well as the Alpha- and Betaproteobacteria phages or prophages. A schematic for predicted lysogeny control through MarR transcriptional regulation is shown in **Figure 42B.** MarR transcriptional regulators are an ancient and widespread family of transcription factors that regulate gene expression in response to stresses (Deochand & Grove, 2017; Pérez-Rueda, Collado-Vides, & Segovia, 2004). The paradigm for MarR transcriptional control is repression through recognition and binding to palindromic DNA sequences which inhibit gene transcription, until binding of a ligand releases MarR (Deochand & Grove, 2017; Grove, 2013; Wilkinson & Grove, 2006). Recent evidence has demonstrated that MarR-like proteins are capable of repressing expression of a number of prophage genes in *Corynebacterium*

glutamicum and *Bacillus subtilis* (De San Eustaquio-Campillo, Cornilleau, Guérin, Carballido-López, & Chastanet, 2017; Hünnefeld, Persicke, Kalinowski, & Frunzke, 2019). Consistent with the model that RNA polymerase acts as a "master switch" for the lytic cycle, examination of the MarR-like protein, integrase, and RNA polymerase locus in phage Pasto revealed two palindromic sequences upstream and downstream of the putative RNA polymerase promoter which may act as MarR binding sites for the transcriptional repression of RNA polymerase (**Figure 42B**). These MarR-like genes are also located in a conserved position directly upstream of the phage integrase, further pointing to a role in the lytic/lysogenic switch. The role of the MarR-like gene involvement in lysogeny maintenance is also supported by RNA-seq of the uninduced Pasto lysogen (**Figure 42B**). RNA-seq of phage lysogens can reveal which genes continue to be expressed while integrated and which may play roles important to lysogen maintenance or accessory proteins that may alter host phenotypes (Owen et al., 2020). RNA-seq revealed a small number of Pasto genes continually expressed from the lysogen. The normalized expression levels in Transcripts per Kilobase Million (TPM) are reported for the genes putatively involved in the lysogeny/lytic switch in **Figure 42B** with host *rpoB* and *secA* as reference. Pasto TPM expression levels of all coding genes are available in **Figure 43**. As predicted, the MarR-like gene remains highly expressed from the lysogen (241 TPM) while the T7-like RNA polymerase and the small hypothetical protein directly upstream remain largely repressed, with 5.58 TPM and 0.78 TPM respectively. Other than the MarR-like gene and integrase, the other two most expressed genes from Pasto lysogen were the small putative membrane protein (1592

TPM) immediately downstream of integrase and the late gene o-spanin (152 TPM). Given the abundance of the small membrane protein and its location, it likely represents a superinfection exclusion mechanism. Reads mapping to the o-spanin only covered a ~115 bp region at the 3' end of the gene, and the function of the small transcript is unknown. Notably, the T7-like cyanophages recently described to integrate but not form stable lysogens (Shitrit et al., 2021) also contain MarR-like repressors in the conserved position. However, utilizing the same strategy as used with Pasto, no putative palindromic MarR binding sites were identified perhaps explaining transient integration without stable lysogen formation. While the putative MarR-like repressor of Pasto and the phage-like repressor of ProddE seem to be the main recognizable mechanisms for lysogeny control in T7-like prophage elements, together they only account for less than half of the lysogeny control mechanisms in the identified prophage elements. This suggests that the remaining prophage elements may have novel, cryptic, or incomplete lysogeny control systems.

Expression of ProddE from the uninduced lysogen was similarly explored to infer mechanisms of lysogen maintenance, and RNA expression levels can be seen in **Figure 44**. Genes believed to be involved in the lysis/lysogeny decision, the CI-like repressor, RNA polymerase, and anti-repressor had TPM values of 61, 14, and 160, respectively. As expected, the RNA polymerase was expressed at a very low level, however a number of other genes putatively believed to be under the control of the phage encoded RNA polymerase, if following T7-like transcription model, were also transcribed to high levels including the endolysin, holin, and excisionase. At first glance,

phage DNA polymerase was also expressed at an extremely high level, however manual examination of the mapped reads to this region revealed most reads specifically mapped to an intron present in the DNA polymerase. Otherwise, coverage in this region approximates the other genes at this locus. Expression of proteins associated with the lytic cycle are presumably due to high levels of spontaneous induction. For this reason, RNAseq methods were inconclusive in determining insights into the ProddE repression strategy as true genes expressed from the lysogen were likely obfuscated by sequencing noise originating from a mixed state of lysogeny/lytic cycles. Also notably, the transcription levels immediately after the major capsid protein dramatically are dramatically reduced, this supports the presence of a strong phage terminator conserved at this site, just as in T7, and the expression of genes upstream of this terminator are consistent with transcription of polycistronic transcripts coding class II and some class III genes in T7 before terminating after T7 major capsid protein (Gene 10) (Carter, Morris, & McAllister, 1981; Jack et al., 2017). Thus, additional experimentation is required to confirm the role of the putative CI-like repressor in the stable maintenance of ProddE lysogeny.

The other major element necessary for the temperate phage lifestyle, the integrase, is more highly conserved within the T7-like temperate phages than the lysogeny maintenance genes. Of the 45 identified prophage elements, 43 encoded recognizable tyrosine recombinase integrases, as did ProddE and Pasto. These 45 integrases share an average of 30.7% identity. Furthermore, 42/45 of the integrases examined contained the Hp1-like integrase conserved domain (cd00796). Those

elements lacking identifiable integrases presumably lost them after their initial integration into the host chromosome and may therefore be defective temperate prophages incapable of excision.

Aside from the lysogeny systems used, there remain many unanswered questions on the infection processes of these phages, which have to date only been studied in the context of lytic infections. Phage T7 is famously dependent on transcription by its own RNA polymerase to complete DNA ejection (Mollineux, 2006), but certain mutations in T7 gp16 (here called internal virion protein D) will allow for transcription-independent DNA ejection (García & Molineux, 1996), and we note that this protein is not well conserved in this group of phages (**Figure 39**). Following DNA ejection, the decision for lysogeny must be rapid and decisive given what is known of the T7 infection cycle, in which the host chromosome is degraded via a phage-encoded endonuclease (gp3) and exonuclease (gp6) (Mollineux, 2006). Both these proteins are conserved among the identified T7-like prophage elements, particularly the gp6 exonuclease (**Figure 40**), so it is assumed the T7-like temperate phages similarly are able to degrade the host genome in their lytic pathways. In well-studied temperate phages like Lambda multiple rounds of DNA replication occur before the decision to enter the lysogenic state can be made (Shao, Trinh, & Zeng, 2019). A single cleavage of the host chromosome could be fatal to the host cell, thus the lysogeny decision in these T7-like temperate phages must preclude the expression of any Class II or Class III genes, including those required for morphogenesis or DNA replication.

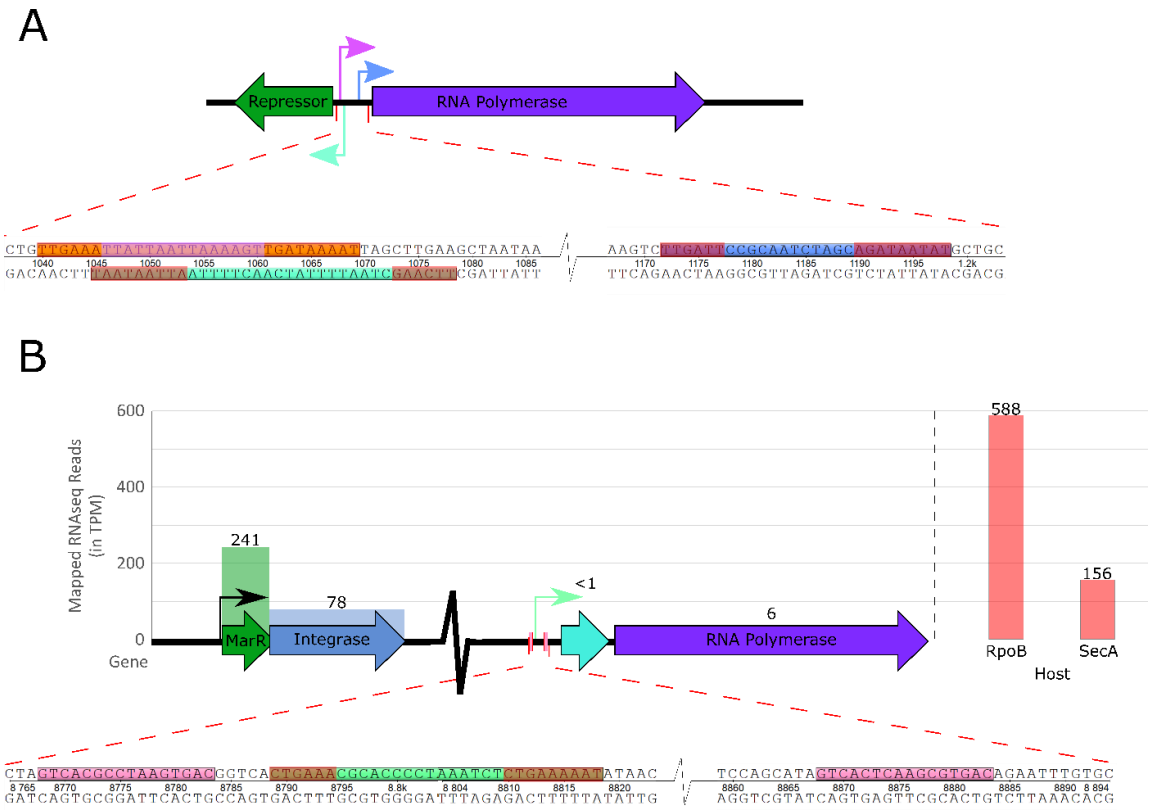


Figure 42. Putative Lysogeny Control Regions of Phages ProddE and Pasto. A) Schematic of ProddE lysis/lysogeny control region, with promoters as predicted by BPROM; -10 and -35 regions are shaded in a contrasting color at the ends of each promoter region. Reminiscent of the Lambda PRM, PRE, and PR arrangement, two predicted promoters (pink/orange and blue/purple) control RNA polymerase transcription, which likely then initiates the ProddE lytic cycle through transcription from ProddE-specific RNA polymerase promoters, while an antisense promoter (teal/brown) controls repressor transcription. B) The predicted lysogeny control region of Pasto. The predicted lysogeny control promoter (green/brown) is flanked by MarR-like protein binding sites (pink boxes). MarR-like transcriptional repressor proteins likely bind flanking palindromic regions of the promoter, suppressing transcription and

controlling RNA polymerase expression. Mapped RNA-seq reads from an uninduced lysogen support the MarR role in lysogeny maintenance. The *marR* gene is expressed from the lysogen at a greater level than the majority of Pasto genes while integrated. Expression levels of host housekeeping genes *rpoB* and *secA* are shown as a reference, (right). RNA-seq read coverage is normalized to transcripts per kilobase million (TPM) and provided above each gene.

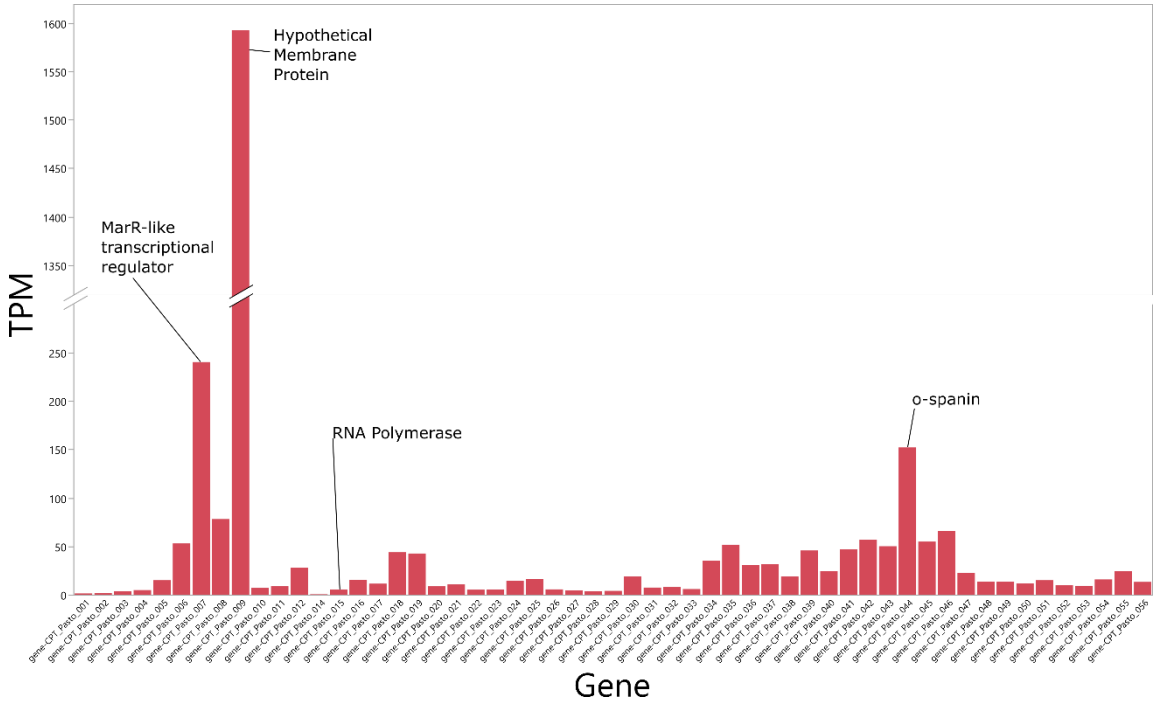


Figure 43. RNA Expression Levels of Phage Pasto.
Mapped RNA-seq reads from an uninduced lysogen over the entirety of Pasto genome. RNA seq reads mapped with FeatureCount. Genes seemingly expressed at high level labeled as well as the RNA polymerase to exemplify its repression. RNA-seq read coverage is normalized to transcripts per kilobase million (TPM) and provided above each gene.

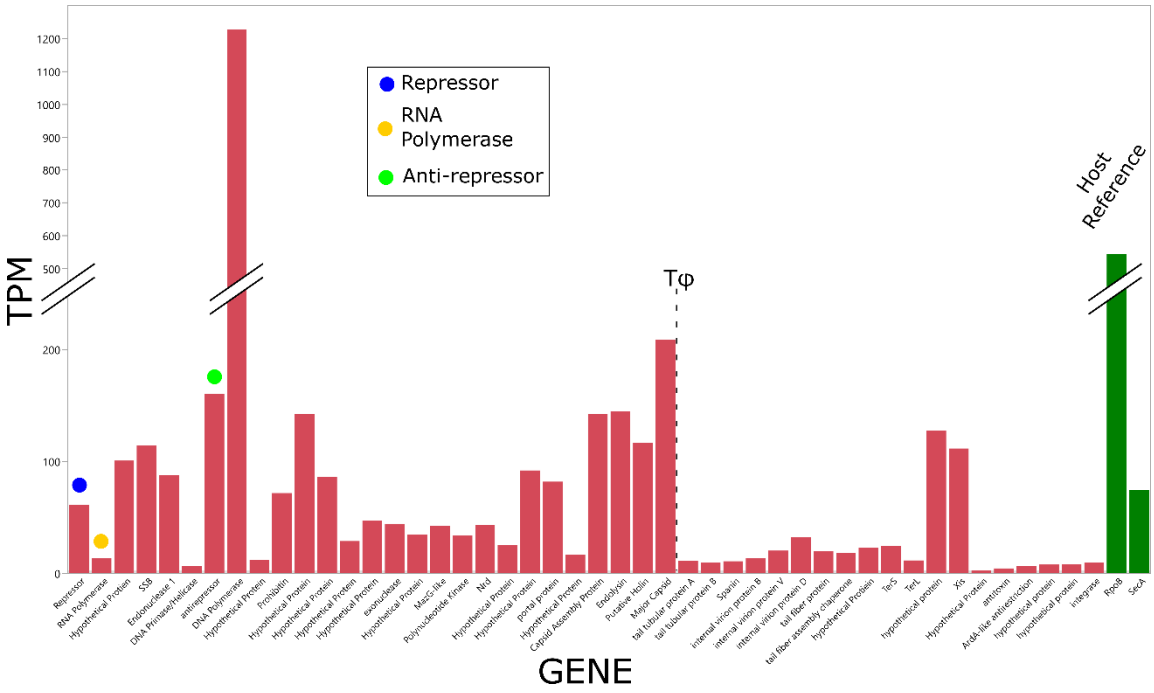


Figure 44. RNA Expression of ProddE.
Mapped RNA-seq reads from an uninduced lysogen over the entirety of ProddE genome. RNA seq reads mapped with FeatureCount. Genes seemingly involved in lytic switch labeled according to legend, Repressor (61 TPM), RNA Polymerase (14 TPM), and Anti-repressor (160 TPM). Host RpoB and SecA provided as point of reference in green. RNA-seq read coverage is normalized to transcripts per kilobase million (TPM). Position of conserved T7 terminator after major capsid (Gene 10 in T7) is denoted by dotted line.

Coevolution of T7-like temperate phages with their hosts

Our data demonstrate that temperate T7-like phages exist in nature, that these elements are widely distributed in Gram-negative bacteria, and that they are remarkably well-conserved in both gene content and gene order, suggesting origination from a

common ancestor. The high levels of sequence divergence observed would indicate that these elements have been diverging for a considerable amount of time. We sought to determine if these elements have been primarily acquired by vertical inheritance in the process of long-term coevolution with their hosts, or if they are more recent arrivals acquired by horizontal gene transfer.

In general, the %G+C content of prophages has been shown to correlate well to the overall %G+C content of the host (Almpanis, Swain, Gatherer, & McEwan, 2018). The %G+C contents of each identified T7-like prophage element, ProddE, and Pasto were plotted against the %G+C of the bacterial host genome in which they were identified (**Figure 45**). Despite the hosts having a large range in %G+C content, ranging from 40.5% to 70.3%, the %G+C content of each T7-like temperate phage element correlates well with its respective host's %G+C, with an R value of .92. The largest discrepancy from this trend was observed in a prophage element identified within a Negativicutes in the genus *Megasphaera,* although this was only a ~10% difference in G+C content. This strong positive correlation implies these prophage elements have resided in their respective hosts for a relatively long time, and suggests a pattern of long-term residence and vertical inheritance of these elements rather than rampant horizontal transfer.

Given the low DNA and protein sequence identity between the T7-like temperate phage elements, classical phylogenetic methods relying on multiple sequence alignments are unreliable. Thus, phylogenetic trees were constructed based on BLASTp pairwise alignments of both the prophage RNA polymerases and the RNA polymerase beta

subunits of the respective hosts (**Figure 46**). As would be expected, the alignment of the host beta subunits perfectly recapitulates the host taxonomic groupings into the Alpha-, Beta-, Gamma-, Deltaproteobacteria, and the Negativicutes (**Figure 46A**). If these T7-like temperate phage elements have exclusively undergone long-term residence and co-evolution with their hosts, the phage RNA polymerase tree should mirror that of the host tree and produce clades that reflect host phylogeny. As can be seen in **Figure 46B**, the phage RNA polymerase clades generally separate by taxonomic class of the host; however, this clustering is not perfect and there are clades containing proteins from multiple host classes. For example, a single clade contains proteins from both gamma- and betaproteobacterial elements (**Fig. 46 box I**) and another contains proteins from both gamma- and alphaproteobacterial hosts (**Fig. 46 box II**), indicating that these elements have likely crossed bacterial class boundaries. The ProddE RNA polymerase, as the only representative from the Deltaproteobacteria, clusters more closely with the alphaproteobacterial members, suggesting a possible origin for this prophage element. It is also noted that the average protein sequence similarity of the host RNA polymerase beta subunits is far greater than that of the phage or prophage RNA polymerases (77% vs. 36%), suggesting a much greater mutation rate in the phage compared to the host. Phages such as Lambda and T4 have been reported to exhibit mutation rates that are on the order of 100-fold greater than their bacterial hosts when replicating lytically (Drake, Charlesworth, Charlesworth, & Crow, 1998); thus, the elevated rate of sequence divergence observed in the phage proteins may represent regular propagation via the phage lytic cycle in addition to being replicated passively with the host chromosome as

prophages, although additional divergence due to varied selection pressures can also not be excluded.

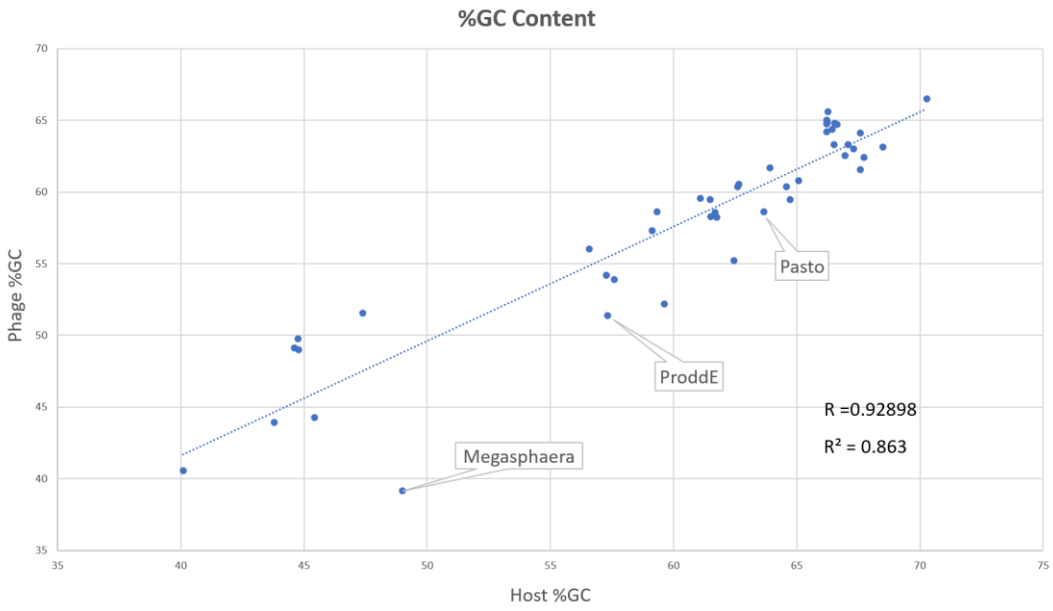


Figure 45. The G+C Content of Prophage Regions Correlates to Host G+C Content. Percent G+C content from the identified prophage elements, as well as ProddE and Pasto, (y-axis) plotted against the overall host %G+C content. Prophage vs Host %G+C shows a positive correlation with an R value of .92898 and an R^2 value of .863. ProddE and Pasto locations are shown as indicated, as well as the prophage element from *Megasphaera* which had the greatest discrepancy between prophage region and host genome %G+C.

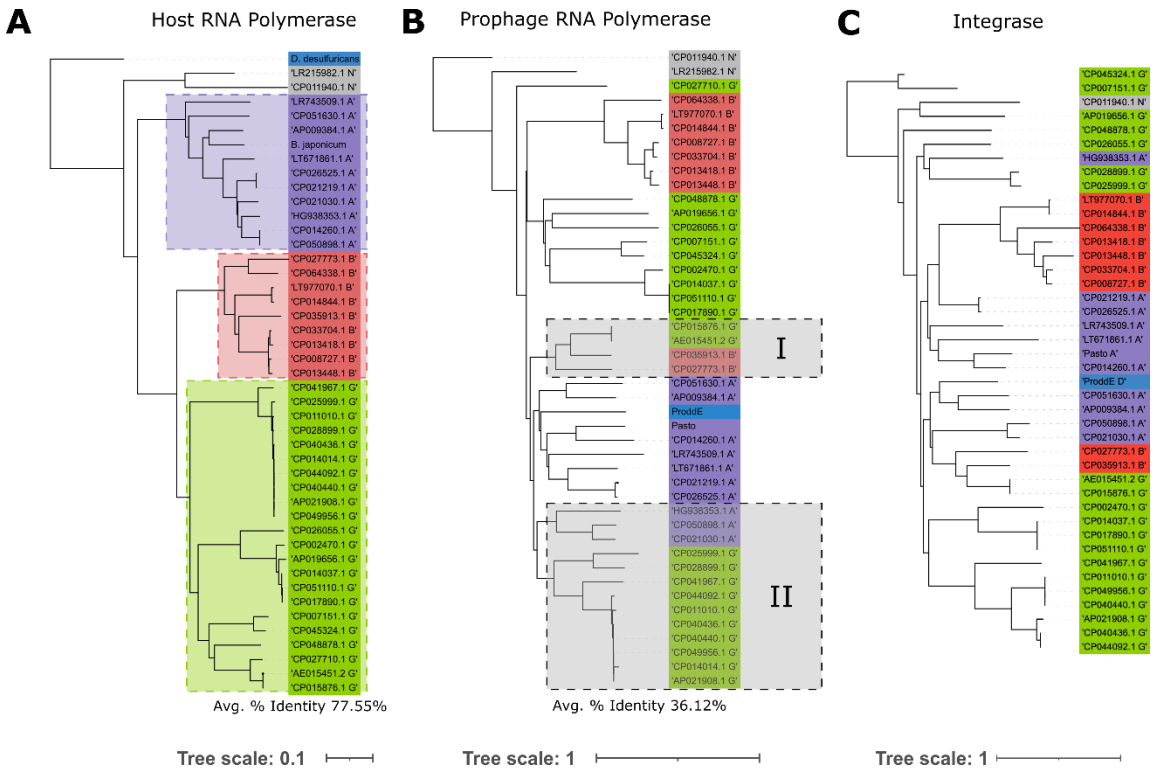


Figure 46. Phylogenetic Trees Comparing Host RNA Polymerase, T7 RNA Polymerases and Integrase.
Phylogenetic trees were generated via pairwise alignment of protein sequences in BLASTp and fastest minimum evolution. A) Tree of RNA polymerase beta subunits from respective prophage hosts. Letters following accession numbers indicate host class: A, B, G, and D are Alpha, Beta, Gamma, and Deltaproteobacteria respectively; N is Negativicutes. Colors and boxes correspond to classes and clades comprised of a single class. Grey=Negativicutes, Green=Gammaproteobacteria, Red=Betaproteobacteria, Purple=Alphaproteobacteria, Blue=Deltaproteobacteria. As a proof of principle, a tree of host RNA polymerases recapitulates the taxonomic classification and monophyletic clades as would be expected by vertical inheritance of a gene. B) Tree of phage RNA polymerases from identified T7-like temperate elements. The phage RNA polymerases

form similar clusters to those observed with bacterial RNA polymerases, but also some polyphyletic clades (boxes I and II) indicating more recent lateral transfers. C.) Integrase, prophages lacking an integrase are omitted. B and C display similar tree topologies, implying similar evolutionary histories and no high levels of acquisition or loss of these elements.

5.3. Conclusion

Based on the evidence provided above, we demonstrate that there is a temperate lineage of T7-like phages capable of lysogen formation based on experimental evidence in this study **(Figures 35 and 36)**, and consistent with observations in other phage-host systems (Martínez-García et al., 2015; Zhao et al., 2019). The identification of numerous integrated T7-like prophage elements, in multiple classes of Gram-negative bacteria**,** with high levels of gene conservation and synteny (**Table 19 and 20, Figures 37 - 40**) points towards a common and ancient origin of this temperate phage lineage. The levels of protein sequence divergence coupled with the strong correlation of prophage %G+C to that of their respective hosts **(Figure 45)** suggests these prophage elements have resided in these hosts for long enough to become adapted, and are not solely the result of more recent horizontal gene transfer events, although such transfers do appear to have occurred in the past as trees of phage and bacterial proteins are not perfectly congruent (**Figure 46**). These functional temperate phages capable of integration, prophage maintenance and induction raise questions about the evolutionary origins of the virulent

Autographiviridae, and whether the temperate lifestyle is ancestral or derived in these phages.

If the ancestral form of these phages was temperate, this would imply that modern virulent T7-like phages could be derived from one or more ancestors that lost their lysogeny-associated functions, rendering them obligately lytic; in this model, conversion to the virulent lifestyle was then followed by additional specialization for occupancy of this new niche. Conversions of phages from temperate to virulent have been observed in phages infecting large dairy fermentation operations in which accelerated evolution can be analyzed (Harald Brüssow & Frank Desiere, 2001; Deveau, Labrie, Chopin, & Moineau, 2006; Søren, Mills, Djordjevic, Israelsen, & Klaenhammer Todd, 2001; Relano, Mata, Bonneau, & Ritzenthaler, 1987). In these cases, virulent *Lactococcus and Streptococcus* phages dominant in dairy environments show evidence of defective or missing genes necessary for lysogeny and high nucleotide identities to known temperate phages suggesting they are directly and recently derived from temperate phages (Søren et al., 2001). Conceptually, the conversion from temperate to virulent by loss of function presents a more parsimonious evolutionary path than the reverse scenario, in which a compatible lysogeny module must be acquired by the virulent phage. Such an acquisition would require, at minimum, an integrase/excisionase pair with cognate *att* sites, and a lysogenic repressor capable of blocking Class II and Class III transcription via one or more compatible operator binding sites. While both paths are possible, they do not appear to be equally probable, and evolution from temperance to virulence likely occurs at a much greater rate than the reverse scenario

due to simple loss of function mutations. If the temperate form of these phages is derived from a virulent ancestor, this would have to be a singular and ancient event to account for the distribution of these elements across multiple bacterial classes (**Table 19 and Figure 41B**), their phylogenetic relationships (**Figure 46**)**,** and the %G+C adaptation (**Figure 45**)**,** which all suggest a significant component of vertical inheritance in the evolution of these elements. If the observed T7-like prophage elements arose from multiple conversions from virulence to temperance, this would require that temperate lifestyle genes (repressor, integrase, etc.) to have been acquired multiple times, resulting in high diversity of these genes. There is some diversity observed in the predicted lysogenic control genes in these phages, which can be divided into three major groups: the CI-like repressor identified in 5/47 elements including ProddE, the MarR-like transcriptional regulators identified in 10/47 elements including Pasto, and unidentified mechanisms (32/47 elements). However, all 45 of the identified integrases of the T7-like temperate elements are tyrosine recombinases, and 42/45 are of the same conserved domain family, the Hp1-like recombinases (cd00796). If integrase genes were acquired independently, we might expect greater variation among the types of integrases acquired, including serine or tyrosine integrases of varying domain classes. The average percent identity of identified integrases, 30.7%, is comparable to that of other conserved homologs like the RNA polymerase (36%) or the tail tubular protein B (33%), indicating these genes have resided within the ancestral T7-like phage for similar amounts of time. Likewise, a phylogenetic tree of integrases constructed using the same methods as

shown in **Figure 46C** produces clusters similar to those observed for the phage RNA polymerases, implying a similar evolutionary history.

Given these observations, we propose a scenario in which a temperate version of a proto-T7-like phage, as exemplified by ProddE or Pasto, was the ancestral state from which the virulent *Autographivirdiae* genera radiated. In this model, an early temperate version of T7-like phages, but with already distinctly recognizable T7-like features and synteny, coevolved with an early Proteobacteria, leading to their presence in diverse genera of Proteobacteria via a combination of vertical inheritance and lateral transfer. Mutations render the lysogeny genes inactive and convert the phage to a virulent lifestyle, which then speciates further toward strict virulence. In this scenario, virulent derivatives could be considered as a sort of "cheater" (Secor Patrick, Dandekar Ajai, & Garsin Danielle) that would enjoy a fitness advantage over the parental temperate phage under some conditions. A potential example of this specialization to a virulent lifestyle may be the presence of seryl-threonyl protein kinase (T7 gene *0.7*) within T7, which is absent from all of the temperate phage elements presented here, implying it has been acquired more recently. The seryl-threonyl protein kinase of T7 is involved shutdown of host RNA polymerase (Severinova & Severinov, 2006) thus increasing efficiency of phage infection, and thereby increasing the virulence of T7.

Autographiviridae phages are currently organized into nine subfamilies and 132 genera, and it remains possible that this family or some of its subfamilies are polyphyletic (Adriaenssens et al., 2020). Thus, it is not clear whether the current known virulent *Autographiviridae* lineages radiated from a single virulent ancestor, whether

they arose multiple times independently, or are the product of recombination events between phage lineages both inside and outside of the *Autographiviridae*. Given the mosaic nature of phage genomes and their propensity to undergo horizontal gene transfer, phages cannot be presumed to follow simple linear paths in their evolution (Lawrence et al., 2002), but rates of gene transfer have been observed to vary widely between phage types, with some exhibiting rampant mosaicism and others forming recognizable monophyletic clades (Mavrich & Hatfull, 2017). The T7-like temperate phages described here appear to fall into the latter category, with these elements showing remarkable conservation in gene content and order (**Figures 39 and 40**). However, these elements are clearly undergoing higher mutation rates than their host bacteria, as shown by the sequence distances observed in phage and host proteins (**Figure 46**), but it is not clear if this is because phage genes have access to a more mutagenic replication pathway via their lytic cycles, or whether this is due to selection pressures that are exerted on the phage but not the host. The high divergence of the T7-like temperate phages is also seen in **Figure 41**, in which the PeA cluster is much more diffuse than the clusters containing virulent *Autographiviridae*. Greater genome variation among temperate phages is not uncommon, when compared to virulent phages infecting the same species (Chopin, Bolotin, Sorokin, Ehrlich, & Chopin, 2001), and it has been proposed that mutations within prophages are more likely to become fixed and persist (Wahl & Pattenden, 2017). Compared to temperate phages, virulent phages do not experience equivalent opportunities for horizontal gene transfer and thus have access to a much more limited gene pool (Mavrich & Hatfull, 2017). This likely holds particularly true for phages like

T7, which rapidly degrades the host chromosome, further reducing chances for recombination (Pajunen, Elizondo, Skurnik, Kieleczawa, & Molineux, 2002). The greater cohesiveness of the virulent phage clusters (e.g., those founded around T7, phiKMV, SP6 and KP34) shown in **Figure 41** may reflect greater conservation due the selective pressures and reduced opportunities for gene transfer imposed by their virulent states, or simply because these groups are the result of more recent divergences that are amplified in number due to sampling biases. These hypothesized conversion events to a strictly virulent lifestyle may have occurred several times, which would account for variations of *Autographiviridae* genome arrangement, as observed in phiKMV or KP34. This is perhaps supported by the additional identification of a number of temperate *Autographiviridae* elements in the phiKMV arrangement (**Figure 38**), although these were excluded from general analysis. Modifications to the search parameters in this study would perhaps identify more proximal integrated temperate ancestors of these phages.

The T7-like prophage elements identified here are remarkably conserved but also relatively rare, appearing in only a small number of the bacterial sequences queried (**Table 19**). This is in stark contrast to the *Autographiviridae* isolated as virulent phages, which make up over 7% of all *Caudovirales* phage genomes deposited in NCBI GenBank. This discrepancy may simply reflect sampling bias in phage isolation, which focuses on hosts in which virulent T7-like phages happen to be highly successful. In the *Studiervirinae,* of which T7 is the founding member, 78% of the phages in NCBI Virus use *Enterobacteriaceae* as host. It is also possible that the four T7 proteins used for the

initial database searches excluded other related prophage elements that are lacking one or more of these proteins or have diverged sufficiently from T7 that they are no longer able to be aligned by BLASTp. It is notable that our analysis did not identify any integrated T7-like prophages within the Enterobacteriaceae itself, despite this being one of the most heavily sequenced bacterial clades (17% of all genomes in the NCBI Microbes database are members of this family). This makes it unlikely that T7 itself or its modern relatives emerged directly from a temperate element in the Enterobacteriaceae, and likely arose in the distant past via a virulent ancestor of T7 that predates the divergence of the Enterobacteriaceae and coevolved with this lineage, or by a more recent invasion of the Enterobacteriaceae as a virulent phage.

Regardless of the precise origins and evolutionary paths of the temperate T7-like phages and their *Autographiviridae* relatives, we show here the existence of a widespread temperate lineage of T7-like phages within the domain Bacteria. Previously, phages encoding their own T7-like RNA polymerase and showing a conserved T7-like genomic arrangement were synonymous with a strictly virulent lifestyle, however with the identification of conserved temperate versions of T7-like phages, this assumption is no longer necessarily valid. This finding is particularly relevant for the selection of potential phages for therapeutic applications, which strongly favors the use of strictly virulent phages (Jason & Paul, 2010). Classification of a phage as a member of a broad group associated with virulence is clearly not sufficient, and the gene content of each candidate phage must be considered, as the barrier between “temperate” and “virulent” lifestyles may be more permeable than is currently appreciated. Aside from the

possession of a tyrosine integrase, however, there are multiple T7-like prophage elements identified here with no other clearly identifiable lysogeny-associated genes. Further exploration of the novel lysis/lysogeny mechanisms present in these T7-like temperate phages could uncover novel mechanisms of genetic control and also aid in phage engineering, providing potent T7-like phages for applied purposes outside the *Enterobacteriaceae*, to which T7 is restricted.

The current model for the origins of the mitochondria in eukaryotic cells is the acquisition of an ancient endosymbiotic bacterium, likely an ancient *Rickettsia*-like alphaproteobacterium, given the gene content of mitochondrial DNA (Gaspari, Larsson, & Gustafsson, 2004; Masters, Stohl, & Clayton, 1987; Roger, Muñoz-Gómez, & Kamikawa, 2017). It has long been known that the mitochondrial RNA polymerase is related to the T7 RNA polymerase (Masters et al., 1987), and the identification of T7-like prophage elements integrated in diverse bacterial hosts, including the Alphaproteobacteria, provides a plausible mechanism for the presence of this gene if the mitochondrial progenitor harbored a T7-like prophage similar to those described here. This model would allow the proto-mitochondrion to possess its T7-like RNA polymerase as a stably integrated chromosomal element, without the need to invoke more complicated scenarios such as coincidental infection by a virulent T7-like phage or a second, separate horizontal gene transfer event.

In summary, based on the evidence presented here, and recently described elsewhere (Martínez-García et al., 2015; Zhao et al., 2019), there exists a functional group of T7-like temperate phages capable of lysogen formation, and these phages do

not appear to be mere aberrant acquisitions of lysogeny genes due to their membership in a widely dispersed and apparently ancient lineage. Although these phages found in different taxonomic classes of bacteria are highly diverged at the sequence level, they are clearly recognizable by their conserved syntenic relationships and are distantly related. Their existence hints at a temperate origin for modern T7-like phages, the *Studiervirinae*, and perhaps the entire family *Autographiviridae*, as well as a paradigm shift away from the T7-like genomic arrangement and RNA polymerase being indicative of a strictly virulent lifestyle.

5.4. Materials and Methods

Isolation and culture of *Desulfovibrio* spp

Strain *Desulfovibrio desulfuricans* strain Edelweiss was isolated from local anoxic creek water in College Station, Tx, USA following guidelines outlined by Postgate as described in Chapter 4. In brief, ~1ml of anoxic creek water was added to an anaerobic Hungate tube filled with pre-reduced Postgate Medium B. Hungate tubes were transferred to a Vinyl COY anaerobic chamber and incubated at 37 °C for one week under the following atmospheric conditions. After one week, tubes containing black Iron Sulfide Precipitate (FeS) tubes were serial diluted with 10-fold dilutions and plated on PMB agar plates amended with Fe at 37 °C anaerobically. Colonies which arose presenting black colonies indicative of H2S production were selected and sub-cultured to ensure clonality. Genus and species level determination was determined based on morphology as determined by phase contrast microscopy and sequencing of the 16S region. After isolation, *Desulfovibrio desulfuricans* Edelweiss was routinely cultured and

maintained on modified Modified Baar's medium for Sulfate Reducers amenable to bacteriophage work as described in Chapter 4. *Agrobacterium* phage Pasto was isolated and cultured as described previously (Manuel et al., 2021). The taxonomic placement of the host, *Agrobacterium* sp. strain AV1, was made to the genus level based on 16S sequence by RDP classifier (Cole et al., 2014).

Sequencing of *Desulfuricans* spp. and Bacteriophage Genomes

Total Bacterial DNA was extracted from a 48hr culture of Edelweiss using a Qiagen DNeasy Blood and Tissue kit following manufacturer's instructions for bacterial culture. Bacteriophage genomic DNA was extracted as previously described modified Promega Wizard DNA cleanup kit protocol, which includes a DNAse/RNAse treatment step to remove host DNA (E. J. Summer, 2009b). Genomic DNA was sequenced on the Illumina MiSeq platform using Illumina MiSeq V300 cycle reagent chemistry by the TAMU Experimental Genomics Core generating pair-ended 150bp. Reads were quality controlled with FastQC (bioinformatics.babraham.ac.uk), FastX Toolkit (hannonlab.cshl.edu), and assembled using SPAdes 3.5.0 in the Center for Phage Technology (CPT) instance of Galaxy. This bacterial genome assembly resulted in 14 contigs greater than 1000 bp with an average 63-fold coverage. Prophage element ProddE did not result in its own contig, but rather integrated into the *D. desulfuricans* genome. The *D. desulfuricans* genome was closed with the addition of Oxford Nanopore long sequencing reads + short Illumina reads using Unicycler (Wick, Judd, Gorrie, & Holt, 2017) and finally ends confirmed by PCR. Sequencing another *Desulfovibrio* phage, phage dd1, which infects strain Edelweiss resulted in contig assembly of the

infecting phage plus a second phage contig that turned out to be the nonintegrated version of ProddE. The ProddE genome was closed by PCR off the ends of the phage genome and subsequent Sanger sequencing. Resulting closed genome was annotated in the CPT Galaxy Web apollo integrated with Six pack, Glimmer 3, Metagene annotator for gene calls; BLAST and Interpro results for putative gene functions; Aragorn and TransTermHP for tRNA and Rho-independent termination sites, respectively. *Desulfovibrio desulfuricans* Edelweiss and phage ProddE sequences are available at accession (CP077745) and accession (MZ666938). Bacteriophage Pasto was isolated and sequenced as previously described (Manuel et al., 2021), available at accession MT708545.1.

RNA Sequencing of Pasto and ProddE Lysogens

A lysogen of phage Pasto was subjected to RNA extraction during exponential phase growth at an approximate OD of 0.3 using an Invitrogen™ TRIzol™ Plus RNA Purification Kit following manufacturer's protocol. To get DNA-free total RNA, on-column DNase treatment was performed using PureLink™ DNase following manufacturer's user guide. Ribosomal RNA depletion of total RNA samples was then performed using Invitrogen™ RiboMinus™ Bacteria 2.0 Transcriptome Isolation Kit following manufacturer's instructions. cDNA was generated from resulting RNA depleted from rRNA with NEBNext® UltraTM II RNA Library kit following manufacturer's instructions. cDNA was then sequenced on Novaseq6000 platform with 1.5v chemistry. Snap frozen cell pellets of *Desulfovibrio desulfuricans* Edelweiss (containing the ProddE prophage) were prepared from two independent rapidly growing

cultures at approx. 0.4 OD had RNA extracted and sequenced by Novogene Co. on a NovaSeq 6000 platform. Reads were quality controlled using FASTQC (bioinformatics.babraham.ac.uk), reads were then first mapped to assembled host genome (non-lysogenized) consisting of concatenated contigs using BowTie2 (Langmead & Salzberg, 2012c) to deplete host RNA reads. Unmapped reads were then mapped to Pasto genome. Resulting read alignments were then counted on a per gene basis using Featurecounts (Liao, Smyth, & Shi, 2013). Raw feature counts of host and Pasto mapped reads were then normalized as Transcripts Per Kilobase Million (TPM) (Conesa et al., 2016). For reference, TPM housekeeping genes *rpoB* and *secA* were also provided (KOCH ET AL., 2019).

Demonstration of ProddE Functionality and Pasto Lysogen Formation

In absence of an appropriate plating host, primers were designed to go off the ends of ProddE genome as integrated into the *D. desulfuricans* Edelweiss. Such primers would only produce a PCR product if the phage genome was capable of excision and the attL and attR sites were brought together. Thermocycler conditions for this reaction were as follows: Initial Denature for 5min at 95C, 30x cycles of (95 ℃ for 30 s, 55 ℃ for 30 s, 72 ℃ for 30 s), final elongation of 72 ℃ for 7min. To demonstrate ProddE packaging into virions and protection from DNAase treatment, short sequencing reads were mapped using BowTie2 at default settings (Langmead & Salzberg, 2012c). Number of mapping reads were compared to the *D. desulfuricans* closed genome (with the ProddE boundaries removed), ProddE, and the phage with which the ProddE genome was assembled. Furthermore, reads from this sequencing run were visualized with BRIG

(Alikhan, Petty, Ben Zakour, & Beatson, 2011) to demonstrate abundance of ProddE reads were not due to contaminating host DNA.

To isolate lysogens and demonstrate Pasto was capable of the lysogenic cycle, high titer stocks of phage Pasto were serially diluted 10-fold and spotted on top agar overlay lawns of *Agrobacterium* sp. strain AV1 and cultured at 28 ℃ on L-arabinose agar at 28 °C. Following incubation, survivors from zones of Pasto clearing were isolated and subcultured 3x to ensure clonality and cultures were free of residual phage. Isolates resistant to Pasto were then cultured and spotted with serially diluted phage Pasto as described above to determine sensitivity and efficiency of plating. Furthermore, to screen for spontaneous induction from putative lysogens, supernatant from centrifuged and filter sterilized O/N culture was spotted and serial diluted onto the non-lysogenic host, as well as supernatant from the non-lysogenized host as a negative control.

Identification of Other T7-like Temperate Phage in the NCBI Database

To identify existence of similar integrated prophage elements in the NCBI database, tblastn was performed on the nr/nt NCBI database, with restricted to Bacteria (taxid:2) using T7's distinctive RNA polymerase (NP_041960.1) as the query. Returned hits from cloning vectors or cloned strains were excluded from analysis. Remaining accession ID's that yielded hits were the basis for a custom blast database. This custom blast database was queried via tblastn for T7 RNA polymerase, DNA Polymerase, Tail Tube B, and TerL. Bacterial genomes missing homologs of these genes were excluded, bacterial accession ID's were further excluded if gene homolog coordinates were not

within a 25 kb to 50 kb locus, ensuring hits represented a single T7-like prophage element. Duplicate prophages from the same species were collapsed. From this list of accessions, integrated prophage element boundary was determined by and extracted by use of PHASTER; those that were not identified by PHASTER were manually extracted.

Comparison of ProddE, Pasto, and Bacteriophage T7, as well as Similar T7-like Temperate Phage

Protein to protein comparison of ProddE putative gene products was performed by blastP with an E value cut off score of 1E-5 to T7 blast database made of proteins from NCBI T7 reference sequence NC_001604.1. Similarly, four representative prophage elements from diverse species, as well as phage Pasto, from bacterial accessions: AP009384.1, CP035913.1, LR215982.1, and CP015876.1 were manually annotated, as ProddE was described above, and compared by BLASTp. BLASTp results of ProddE and T7 percent ID used to construct the genome comparison map in **Figure 37 and Figure 39**.

vContact2 Taxonomy and Viral Clustering

Putative phage boundaries were extracted from the bacterial genomes as described above, along with ProddE and Rhizobium phage Pasto, and used as input for vContact2 v0.9.8 using the RefSeqv88 database (Ho Bin Jang et al., 2019; Bolduc et al., 2017) on the CyVerse platform (Merchant et al., 2016). Prophage gene prediction was performed by Prodigal (Hyatt et al., 2010). Resulting gene calls were used as input for vContact using default settings and blastP option. Resulting network file output was

visualized in Cytoscape (Shannon et al., 2003). RefSeq outliers hidden from view to assist in visualization, and nodes color coded according to figure legend.

RNA Polymerase Phylogenetic Tree Comparison

RNA polymerase amino acid sequence from prophage elements, ProddE, and Pasto were extracted. Those RNA polymerases that contained an intron were manually concatenated. RNA polymerase subunit beta was extracted from respective bacterial host genomes that contained T7-like temperate prophages. The host of phage Pasto has no sequence for its direct host, so a RNA polymerase beta subunit representative of the *Rhizobium/Agrobacterium* group was used. Resulting host and phage sequences were compared via BLASTp to form pairwise alignments and initial tree produced using the integrated Blast tree view with Fast Minimum evolution, Max Seq Difference 0.85, and following Grishin General (Protein) distance model. Resulting trees were then annotated in the Interactive Tree of Life tool (iTol)(Letunic & Bork, 2021).

6. CONCLUSIONS

In conclusion, the gut microbiome plays an integral part of human health and disease, particularly chronic inflammatory diseases of the intestine. As our understanding of microbiome and its altered states continues to grow, our ability to use this knowledge and selectively engineer the microbiome will become increasingly important. Phages may represent a unique tool in this regard given their natural modulation of bacterial populations and narrow host ranges. In this work, efforts to model environmental enteric disorder, EED, did lead to a detectable dysbiosis in both the piglets receiving acute and chronic doses of pathogenic *E. coli*. However, symptoms classically associated with EED, such blunting of intestinal villi, were not achieved, as such the model presented here may be more representative of a general enteropathy. While the phages isolated in an effort to alleviate this modeled enteropathy were inconclusive in terms of their efficacy to reduce the *E. coli* levels and induced clinical signs, two represent part of a recently described novel group of phages with optimum lytic temperatures below 37 °C. While not useful for therapeutic applications, "low-temperature" phages may be effective biocontrol agents in food safety applications, in which food is maintained at ambient or below temperatures like during the transport and storage of meat and produce for retail.

Next, phages were isolated against a *Desulfovibrio desulfuricans* in an effort to evaluate the potential of phage for the control of *Desulfovibrio* spp. in the gut as little to no modern examinations of *Desulfovibrio* phage exist. Phages isolated were all shown to be highly novel compared to existing phages isolated, likely justifying 3 novel genera of

phage. Although the majority of phage isolated are predicted to be temperate, they nevertheless displayed ability to reduce *Desulfovibrio* growth over several days supporting their use as a novel tool for reduction of *Desulfovibrio*.

Finally, during the exploration of the *Desulfovibrio* prophages, the identification of an active novel prophage element from within the *Desulfovibrio* host led to the identification of a widely distributed novel group of T7-like temperate members of the *Autographiviridae* with novel mechanisms of lysogen maintenance. Furthermore, the existence of a number of these T7-like temperate elements suggests a temperate origin for the *Autographiviridae* and disrupts the paradigm of strictly virulent T7-like phages.

A number of the concepts identified in this work warrant further study. For example, further exploration of the low-temperature phages and the mechanisms which enhance their performance at lower temperatures may lead to enhanced formulations of phage-based biocontrol products. The continued exploration and isolation of *Desulfovibrio* phages may uncover interesting aspects of phage-host biology absent from classic model bacterial species and further support their use to control *Desulfovibrio* in clinical or industrial settings. Further explorations of the T7-like temperate phages and their potential engineering may supply potent T7-like phages for the control of a broader range of bacterial species than the *Enterobacteriaceae*, which most T7-like phage identified are currently limited to.

REFERENCES

Abedon, S. T. (2000). The murky origin of Snow White and her T-even dwarfs. *Genetics, 155*(2), 481-486.

Adams, M. H. (1959). *Bacteriophages*. New York: Interscience Publishers.

Adriaenssens, E. M., Sullivan, M. B., Knezevic, P., van Zyl, L. J., Sarkar, B. L., Dutilh, B. E., . . . Krupovic, M. (2020). Taxonomy of prokaryotic viruses: 2018-2019update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Archives of Virology, 165*(5), 1253-1260. doi:10.1007/s00705-020-04577-8

Aggarwala, V., Liang, G., & Bushman, F. D. (2017). Viral communities of the human gut: metagenomic analysis of composition and dynamics. *Mobile DNA, 8*(1), 12. doi:10.1186/s13100-017-0095-y

Aksyuk, A. A., & Rossmann, M. G. (2011). Bacteriophage assembly. *Viruses, 3*(3), 172-203. doi:10.3390/v3030172

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics, 12*(1), 402. doi:10.1186/1471-2164-12-402

Almpanis, A., Swain, M., Gatherer, D., & McEwan, N. (2018). Correlation between bacterial G+C content, genome size and the G+C content of associated plasmids and bacteriophages. *Microbial genomics, 4*(4), e000168. doi:10.1099/mgen.0.000168

Arango Duque, G., & Descoteaux, A. (2014). Macrophage Cytokines: Involvement in Immunity and Infectious Diseases. *Frontiers in Immunology, 5*(491). doi:10.3389/fimmu.2014.00491

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res, 44*(W1), W16-21. doi:10.1093/nar/gkw387

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D. R., . . . Meta, H. I. T. C. (2011). Enterotypes of the human gut microbiome. *Nature, 473*(7346), 174-180. doi:10.1038/nature09944

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., . . . Honda, K. (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature, 500*(7461), 232-236. doi:10.1038/nature12331

Atkison, C. L., Boeckman, J., Newkirk, H., Liu, M., Gill, J. J., Cahill, J., & Ramsey, J. (2019). Complete Genome Sequence of Escherichia coli Myophage Mangalitsa. *Microbiol Resour Announc, 8*(38), e01045-01019. doi:10.1128/MRA.01045-19

Avrani, S., Schwartz, D. A., & Lindell, D. (2012). Virus-host swinging party in the oceans: Incorporating biological complexity into paradigms of antagonistic coexistence. *Mobile genetic elements, 2*(2), 88-95. doi:10.4161/mge.20031

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., . . . Wang, J. (2015). Dynamics and Stabilization of the Human Gut Microbiome

during the First Year of Life. *Cell Host & Microbe, 17*(5), 690-703. doi:https://doi.org/10.1016/j.chom.2015.04.004

Banerjee, P., Carmelo, V. A. O., & Kadarmideen, H. N. (2020). Integrative Analysis of Metabolomic and Transcriptomic Profiles Uncovers Biological Pathways of Feed Efficiency in Pigs. *Metabolites, 10*(7), 275. doi:10.3390/metabo10070275

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology, 19*(5), 455-477. doi:10.1089/cmb.2012.0021

Barbosa, J. A., Rodrigues, L. A., Columbus, D. A., Aguirre, J. C. P., Harding, J. C. S., Cantarelli, V. S., & Costa, M. d. O. (2021). Experimental infectious challenge in pigs leads to elevated fecal calprotectin levels following colitis, but not enteritis. *Porcine Health Management, 7*(1), 48. doi:10.1186/s40813-021-00228-9

Bartlett, J. G., Onderdonk, A. B., Cisneros, R. L., & Kasper, D. L. (1977). Clindamycin-Associated Colitis Due to a Toxin-Producing Species of Clostridium in Hamsters. *The Journal of Infectious Diseases, 136*(5), 701-705. doi:10.1093/infdis/136.5.701

Barton, L. L., & Fauque, G. D. (2009). Chapter 2 Biochemistry, Physiology and Biotechnology of Sulfate-Reducing Bacteria. In (pp. 41-98).

Battaglioli, E. J., Hale, V. L., Chen, J., Jeraldo, P., Ruiz-Mojica, C., Schmidt, B. A., . . . Kashyap, P. C. (2018). Clostridioides difficile uses amino acids associated with

gut microbial dysbiosis in a subset of patients with diarrhea. *Science Translational Medicine, 10*(464), eaam7019. doi:10.1126/scitranslmed.aam7019

Belogurov, A. A., Delver, E. P., & Rodzevich, O. V. (1993). Plasmid pKM101 encodes two nonhomologous antirestriction proteins (ArdA and ArdB) whose expression is controlled by homologous regulatory sequences. *Journal of Bacteriology, 175*(15), 4843-4850. doi:10.1128/jb.175.15.4843-4850.1993

Berberov, E. M., Zhou, Y., Francis, D. H., Scott, M. A., Kachman, S. D., & Moxley, R. A. (2004). Relative Importance of Heat-Labile Enterotoxin in the Causation of Severe Diarrheal Disease in the Gnotobiotic Piglet Model by a Strain of Enterotoxigenic Escherichia coli That Produces Multiple Enterotoxins. *Infection and Immunity*, *72*(7), 3914-3924. doi:10.1128/IAI.72.7.3914-3924.2004

Bhandari, S. K., Xu, B., Nyachoti, C. M., Giesting, D. W., & Krause, D. O. (2008). Evaluation of alternatives to antibiotics using an Escherichia coli K88+ model of piglet diarrhea: Effects on gut microbial ecology1. *Journal of Animal Science, 86*(4), 836-847. doi:10.2527/jas.2006-822

Bhardwaj, G., Ko, K. D., Hong, Y., Zhang, Z., Ho, N. L., Chintapalli, S. V., . . . van Rossum, D. B. (2012). PHYRN: A Robust Method for Phylogenetic Analysis of Highly Divergent Sequences. *PLOS ONE, 7*(4), e34261. doi:10.1371/journal.pone.0034261

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., . . . Sullivan, M. B. (2019). Taxonomic assignment of uncultivated prokaryotic virus

genomes is enabled by gene-sharing networks. *Nat Biotechnol, 37*(6), 632-639. doi:10.1038/s41587-019-0100-8

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., . . . Sullivan, M. B. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology, 37*(6), 632-639. doi:10.1038/s41587-019-0100-8

Bjerrum, J. T., Wang, Y., Hao, F., Coskun, M., Ludwig, C., Gunther, U., & Nielsen, O. H. (2015). Metabonomics of human fecal extracts characterize ulcerative colitis, Crohn's disease and healthy individuals. *Metabolomics, 11*, 122-133. doi:10.1007/s11306-014-0677-3

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., . . . Rivera, J. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet, 371*(9608), 243-260. doi:https://doi.org/10.1016/S0140-6736(07)61690-0

Blanton Laura, V., Charbonneau Mark, R., Salih, T., Barratt Michael, J., Venkatesh, S., Ilkaveya, O., . . . Gordon Jeffrey, I. (2016). Gut bacteria that prevent growth impairments transmitted by microbiota from malnourished children. *Science, 351*(6275), aad3311. doi:10.1126/science.aad3311

Blick, A. K., Giaretta, P. R., Sprayberry, S., Bush-Vadala, C., Paulk, C. B., Boeckman, J., . . . Rech, R. R. (2019). Comparison of 2 fixatives in the porcine colon for in situ microbiota studies. *Journal of Animal Science, 97*(12), 4803-4809. doi:10.1093/jas/skz325

Bobay, L. M., Rocha, E. P., & Touchon, M. (2013). The adaptation of temperate bacteriophages to their host genomes. *Mol Biol Evol, 30*(4), 737-751. doi:10.1093/molbev/mss279

Boeckman, J. X., Sprayberry, S., Korn, A., Suchodolski, J. S., Paulk, C., Genovese, K., . . . Gill, J. J. (2021). Development and Characterization of a Weaned Pig Model of Shiga Toxin–Producing E. coli-Induced Gastrointestinal Disease. *bioRxiv*, 2021.2008.2026.457881. doi:10.1101/2021.08.26.457881

Bolduc, B., Jang, H. B., Doulcier, G., You, Z.-Q., Roux, S., & Sullivan, M. B. (2017). vConTACT: an iVirus tool to classify double-stranded DNA viruses that infect Archaea and Bacteria. *PeerJ, 5*, e3243-e3243. doi:10.7717/peerj.3243

Boon, M., Holtappels, D., Lood, C., van Noort, V., & Lavigne, R. (2020). Host Range Expansion of Pseudomonas Virus LUZ7 Is Driven by a Conserved Tail Fiber Mutation. *PHAGE, 1*(2), 87-90. doi:10.1089/phage.2020.0006

Britton, G. J., Contijoch, E. J., Mogno, I., Vennaro, O. H., Llewellyn, S. R., Ng, R., . . . Faith, J. J. (2019). Microbiotas from Humans with Inflammatory Bowel Disease Alter the Balance of Gut Th17 and RORγt+ Regulatory T Cells and Exacerbate Colitis in Mice. *Immunity, 50*(1), 212-224.e214. doi:https://doi.org/10.1016/j.immuni.2018.12.015

Broussard, Gregory W., Oldfield, Lauren M., Villanueva, Valerie M., Lunt, Bryce L., Shine, Emilee E., & Hatfull, Graham F. (2013). Integration-Dependent Bacteriophage Immunity Provides Insights into the Evolution of Genetic

Switches. *Molecular Cell, 49*(2), 237-248. doi:https://doi.org/10.1016/j.molcel.2012.11.012

Brown, E. M., Wlodarska, M., Willing, B. P., Vonaesch, P., Han, J., Reynolds, L. A., . . . Finlay, B. B. (2015). Diet and specific microbial exposure trigger features of environmental enteropathy in a novel murine model. *Nature communications, 6*(1), 7806. doi:10.1038/ncomms8806

Brown, E. M., Wlodarska, M., Willing, B. P., Vonaesch, P., Han, J., Reynolds, L. A., . . . Finlay, B. B. (2015). Diet and specific microbial exposure trigger features of environmental enteropathy in a novel murine model. *Nat Commun, 6*, 7806. doi:10.1038/ncomms8806

Brown Kevin, A., Khanafer, N., Daneman, N., & Fisman David, N. (2013). Meta-Analysis of Antibiotics and the Risk of Community-Associated Clostridium difficile Infection. *Antimicrobial Agents and Chemotherapy, 57*(5), 2326-2332. doi:10.1128/AAC.02176-12

Brüssow, H., & Desiere, F. (2001). Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. *Molecular Microbiology, 39*(2), 213-223. doi:https://doi.org/10.1046/j.1365-2958.2001.02228.x

Brüssow, H., & Desiere, F. (2001). Comparative phage genomics and the evolution of Siphoviridae: insights from dairy phages. *Mol Microbiol, 39*(2), 213-222. doi:10.1046/j.1365-2958.2001.02228.x

Brüssow, H., & Hendrix, R. W. (2002). Phage Genomics: Small Is Beautiful. *Cell, 108*(1), 13-16. doi:10.1016/S0092-8674(01)00637-7

Bucker, R., Schulz, E., Gunzel, D., Bojarski, C., Lee, I. F., John, L. J., . . . Schulzke, J. D. (2014). alpha-Haemolysin of Escherichia coli in IBD: a potentiator of inflammatory activity in the colon. *Gut, 63*(12), 1893-1901. doi:10.1136/gutjnl-2013-306099

Buffie Charlie, G., Jarchum, I., Equinda, M., Lipuma, L., Gobourne, A., Viale, A., . . . McCormick, B. A. (2012). Profound Alterations of Intestinal Microbiota following a Single Dose of Clindamycin Results in Sustained Susceptibility to Clostridium difficile-Induced Colitis. *Infection and Immunity, 80*(1), 62-73. doi:10.1128/IAI.05496-11

Byersdorfer, C. A. (2014). The role of Fatty Acid oxidation in the metabolic reprograming of activated t-cells. *Front Immunol, 5*, 641. doi:10.3389/fimmu.2014.00641

Caffalette, C. A., & Zimmer, J. (2021). Cryo-EM structure of the full-length WzmWzt ABC transporter required for lipid-linked O antigen transport. *Proceedings of the National Academy of Sciences, 118*(1), e2016144118. doi:10.1073/pnas.2016144118

Cahenzli, J., Köller, Y., Wyss, M., Geuking, Markus B., & McCoy, Kathy D. (2013). Intestinal Microbial Diversity during Early-Life Colonization Shapes Long-Term IgE Levels. *Cell Host & Microbe, 14*(5), 559-570. doi:https://doi.org/10.1016/j.chom.2013.10.004

Cahill, J., & Young, R. (2019). Phage Lysis: Multiple Genes for Multiple Barriers. *Advances in virus research, 103*, 33-70. doi:10.1016/bs.aivir.2018.09.003

Cai, Y., Martínez, Y., Li Qiang Yao, Y., & Yu, L. (2016). *Review of enterotoxigenic Escherichia coli enterotoxins* (Vol. 4). Acad. J. Microbiol.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics, 10*(1), 421. doi:10.1186/1471-2105-10-421

Carmelo, V. A. O., Banerjee, P., da Silva Diniz, W. J., & Kadarmideen, H. N. (2020). Metabolomic networks and pathways associated with feed efficiency and related-traits in Duroc and Landrace pigs. *Scientific Reports, 10*(1), 255. doi:10.1038/s41598-019-57182-4

Carter, A. D., Morris, C. E., & McAllister, W. T. (1981). Revised transcription map of the late region of bacteriophage T7 DNA. *Journal of Virology, 37*(2), 636-642. doi:10.1128/jvi.37.2.636-642.1981

Carter, C. D., Parks, A., Abuladze, T., Li, M., Woolston, J., Magnone, J., . . . Sulakvelidze, A. (2012). Bacteriophage cocktail significantly reduces Escherichia coli O157: H7 contamination of lettuce and beef, but does not protect against recontamination. *Bacteriophage, 2*(3), 178-185. doi:10.4161/bact.22825

Casewell, M., Friis, C., Marco, E., McMullin, P., & Phillips, I. (2003). The European ban on growth-promoting antibiotics and emerging consequences for human and animal health. *J Antimicrob Chemother, 52*(2), 159-161. doi:10.1093/jac/dkg313

Casjens, S. (2003). Prophages and bacterial genomics: what have we learned so far? *Molecular Microbiology, 49*(2), 277-300. doi:https://doi.org/10.1046/j.1365-2958.2003.03580.x

Casjens, S. R., & Thuman-Commike, P. A. (2011). Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. *Virology, 411*(2), 393-415. doi:https://doi.org/10.1016/j.virol.2010.12.046

Cha, S. B., Yoo, A. N., Lee, W. J., Shin, M. K., Jung, M. H., Shin, S. W., . . . Yoo, H. S. (2012). Effect of Bacteriophage in Enterotoxigenic Escherichia coli (ETEC) Infected Pigs. *Journal of Veterinary Medical Science, 74*(8), 1037-1039. doi:10.1292/jvms.11-0556

Chang, H.-W., McNulty, N. P., Hibberd, M. C., O'Donnell, D., Cheng, J., Lombard, V., . . . Gordon, J. I. (2021). Gut microbiome contributions to altered metabolism in a pig model of undernutrition. *Proc Natl Acad Sci U S A, 118*(21), e2024446118. doi:10.1073/pnas.2024446118

Chehoud, C., Dryga, A., Hwang, Y., Nagy-Szakal, D., Hollister Emily, B., Luna Ruth, A., . . . Steward Grieg, F. (2016). Transfer of Viral Communities between Human Individuals during Fecal Microbiota Transplantation. *mBio, 7*(2), e00322-00316. doi:10.1128/mBio.00322-16

Chen, L., Xu, Y., Chen, X., Fang, C., Zhao, L., & Chen, F. (2017). The Maturing Development of Gut Microbiota in Commercial Piglets during the Weaning Transition. *Front Microbiol, 8*, 1688. doi:10.3389/fmicb.2017.01688

Chen, L., Zheng, D., Liu, B., Yang, J., & Jin, Q. (2016). VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Research, 44*(Database issue), D694-D697. doi:10.1093/nar/gkv1239

Chen, Y. R., Zhou, L. Z., Fang, S. T., Long, H. Y., Chen, J. Y., & Zhang, G. X. (2019). Isolation of Desulfovibrio spp. from human gut microbiota using a next-generation sequencing directed culture method. *Lett Appl Microbiol, 68*(6), 553-561. doi:10.1111/lam.13149

Chopin, A., Bolotin, A., Sorokin, A., Ehrlich, S. D., & Chopin, M.-C. (2001). Analysis of six prophages in Lactococcus lactis IL1403: different genetic structure of temperate and virulent phage populations. *Nucleic Acids Res, 29*(3), 644-651. doi:10.1093/nar/29.3.644

Clokie, M. R., Millard, A. D., Letarov, A. V., & Heaphy, S. (2011). Phages in nature. *Bacteriophage, 1*(1), 31-45. doi:10.4161/bact.1.1.14942

Clooney, A. G., Sutton, T. D. S., Shkoporov, A. N., Holohan, R. K., Daly, K. M., O'Regan, O., . . . Hill, C. (2019). Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. *Cell Host & Microbe, 26*(6), 764-778.e765. doi:https://doi.org/10.1016/j.chom.2019.10.009

Cock, P. J. A., Grüning, B. A., Paszkiewicz, K., & Pritchard, L. (2013). Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. *PeerJ, 1*, e167. doi:10.7717/peerj.167

Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., . . . Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res, 42*(Database issue), D633-D642. doi:10.1093/nar/gkt1244

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., . . . Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, *17*(1), 13. doi:10.1186/s13059-016-0881-8

Costa, F., Mumolo, M. G., Bellini, M., Romano, M. R., Ceccarelli, L., Arpe, P., . . . Maltinti, G. (2003). Role of faecal calprotectin as non-invasive marker of intestinal inflammation. *Dig Liver Dis*, *35*(9), 642-647.

Costea, P. I., Hildebrand, F., Arumugam, M., Bäckhed, F., Blaser, M. J., Bushman, F. D., . . . Bork, P. (2018). Enterotypes in the landscape of gut microbial community composition. *Nat Microbiol*, *3*(1), 8-16. doi:10.1038/s41564-017-0072-8

Council, N. R. (2012). *Nutrient Requirements of Swine: Eleventh Revised Edition*. Washington, DC: The National Academies Press.

Coutinho, C. M. L. M., Coutinho-Silva, R., Zinkevich, V., Pearce, C. B., Ojcius, D. M., & Beech, I. (2017a). Sulphate-reducing bacteria from ulcerative colitis patients induce apoptosis of gastrointestinal epithelial cells. *Microbial Pathogenesis*, *112*, 126-134. doi:https://doi.org/10.1016/j.micpath.2017.09.054

Coutinho, C. M. L. M., Coutinho-Silva, R., Zinkevich, V., Pearce, C. B., Ojcius, D. M., & Beech, I. (2017b). Sulphate-reducing bacteria from ulcerative colitis patients induce apoptosis of gastrointestinal epithelial cells. *Microbial Pathogenesis*, *112*, 126-134. doi:https://doi.org/10.1016/j.micpath.2017.09.054

Cowley, L. A., Low, A. S., Pickard, D., Boinett, C. J., Dallman, T. J., Day, M., . . . Cain, A. K. (2018). Transposon Insertion Sequencing Elucidates Novel Gene

Involvement in Susceptibility and Resistance to Phages T4 and T7 in Escherichia coli O157. *MBio, 9*(4), e00705-00718. doi:10.1128/mBio.00705-18

Crispim, J. S., Dias, R. S., Laguardia, C. N., Araujo, L. C., da Silva, J. D., Vidigal, P. M. P., . . . de Paula, S. O. (2019). Desulfovibrio alaskensis prophages and their possible involvement in the horizontal transfer of genes by outer membrane vesicles. *Gene, 703*, 50-57. doi:10.1016/j.gene.2019.04.016

Crispim, J. S., Dias, R. S., Vidigal, P. M. P., de Sousa, M. P., da Silva, C. C., Santana, M. F., & de Paula, S. O. (2018). Screening and characterization of prophages in Desulfovibrio genomes. *Sci Rep, 8*(1), 9273. doi:10.1038/s41598-018-27423-z

Cumby, N., Davidson, A. R., & Maxwell, K. L. (2012). The moron comes of age. *Bacteriophage, 2*(4), 225-228. doi:10.4161/bact.23146

Cuthbertson, L., Powers, J., & Whitfield, C. (2005). The C-terminal Domain of the Nucleotide-binding Domain Protein Wzt Determines Substrate Specificity in the ATP-binding Cassette Transporter for the Lipopolysaccharide O-antigens in Escherichia coli Serotypes O8 and O9a *<sup>

</sup>. *Journal of Biological Chemistry, 280*(34), 30310-30319. doi:10.1074/jbc.M504371200

Darling, A. E., Mau, B., & Perna, N. T. (2010). progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. *PLoS One, 5*(6), e11147. doi:10.1371/journal.pone.0011147

Das, S., Jayaratne, R., & Barrett, K. E. (2018). The Role of Ion Transporters in the Pathophysiology of Infectious Diarrhea. *Cellular and molecular gastroenterology and hepatology, 6*(1), 33-45. doi:10.1016/j.jcmgh.2018.02.009

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., . . . Turnbaugh, P. J. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature, 505*(7484), 559-563. doi:10.1038/nature12820

Dawes, J. (1975). Characterisation of the bacteriophage T4 receptor site. *Nature, 256*(5513), 127-128. doi:10.1038/256127a0

de Onis, M., Blössner, M., & Borghi, E. (2012). Prevalence and trends of stunting among pre-school children, 1990–2020. *Public Health Nutrition, 15*(1), 142-148. doi:10.1017/S1368980011001315

De San Eustaquio-Campillo, A., Cornilleau, C., Guérin, C., Carballido-López, R., & Chastanet, A. (2017). PamR, a new MarR-like regulator affecting prophages and metabolic genes expression in Bacillus subtilis. *PLOS ONE, 12*(12), e0189694. doi:10.1371/journal.pone.0189694

DeBoer, M. D., Lima, A. A. M., Oría, R. B., Scharf, R. J., Moore, S. R., Luna, M. A., & Guerrant, R. L. (2012). Early childhood growth failure and the developmental origins of adult disease: Do enteric infections and malnutrition increase risk for the metabolic syndrome? *Nutrition reviews, 70*(11), 642-653. doi:10.1111/j.1753-4887.2012.00543.x

Dejea, C. M., Fathi, P., Craig, J. M., Boleij, A., Taddese, R., Geis, A. L., . . . Sears, C. L. (2018). Patients with familial adenomatous polyposis harbor colonic biofilms

containing tumorigenic bacteria. *Science, 359*(6375), 592-597. doi:doi:10.1126/science.aah3648

DELBRUCK, M. (1946). BACTERIAL VIRUSES OR BACTERIOPHAGES. *Biological Reviews, 21*(1), 30-40. doi:https://doi.org/10.1111/j.1469-185X.1946.tb00451.x

Demerec, M., & Fano, U. (1945). Bacteriophage-Resistant Mutants in Escherichia Coli. *Genetics, 30*(2), 119-136.

Deochand, D. K., & Grove, A. (2017). MarR family transcription factors: dynamic variations on a common scaffold. *Critical Reviews in Biochemistry and Molecular Biology, 52*(6), 595-613. doi:10.1080/10409238.2017.1344612

Desai, C., Handley, S. A., Rodgers, R., Rodriguez, C., Ordiz, M. I., Manary, M. J., & Holtz, L. R. (2020). Growth velocity in children with Environmental Enteric Dysfunction is associated with specific bacterial and viral taxa of the gastrointestinal tract in Malawian children. *PLOS Neglected Tropical Diseases, 14*(6), e0008387. doi:10.1371/journal.pntd.0008387

Dethlefsen, L., & Relman, D. A. (2011). Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences, 108*(Supplement 1), 4554. doi:10.1073/pnas.1000087107

Deveau, H., Labrie, S. J., Chopin, M.-C., & Moineau, S. (2006). Biodiversity and classification of lactococcal phages. *Applied and Environmental Microbiology, 72*(6), 4338-4346. doi:10.1128/AEM.02517-05

Ding, Z.-H., Xu, X.-P., Wang, T.-R., Liang, X., Ran, Z.-H., & Lu, H. (2021). The prevalence of Helicobacter pylori infection in inflammatory bowel disease in China: A case-control study. *PLoS One, 16*(3), e0248427. doi:10.1371/journal.pone.0248427

Dobbins, A. T., George, M., Basham, D. A., Ford, M. E., Houtz, J. M., Pedulla, M. L., . . . Hendrix, R. W. (2004). Complete Genomic Sequence of the Virulent Salmonella Bacteriophage SP6. *Journal of Bacteriology, 186*(7), 1933. doi:10.1128/JB.186.7.1933-1944.2004

Dou, S., Gadonna-Widehem, P., Rome, V., Hamoudi, D., Rhazi, L., Lakhal, L., . . . Abdennebi-Najar, L. (2017). Characterisation of Early-Life Fecal Microbiota in Susceptible and Healthy Pigs to Post-Weaning Diarrhoea. *PLoS One, 12*(1), e0169851. doi:10.1371/journal.pone.0169851

Dowah, A. S. A., & Clokie, M. R. J. (2018). Review of the nature, diversity and structure of bacteriophage receptor binding proteins that target Gram-positive bacteria. *Biophys Rev, 10*(2), 535-542. doi:10.1007/s12551-017-0382-3

Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics, 148*(4), 1667-1686.

Draper, L. A., Ryan, F. J., Smith, M. K., Jalanka, J., Mattila, E., Arkkila, P. A., . . . Hill, C. (2018). Long-term colonisation with donor bacteriophages following successful faecal microbial transplantation. *Microbiome, 6*(1), 220. doi:10.1186/s40168-018-0598-x

Drewes, J. L., White, J. R., Dejea, C. M., Fathi, P., Iyadorai, T., Vadivelu, J., . . . Sears, C. L. (2017). High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *NPJ Biofilms Microbiomes, 3*, 34. doi:10.1038/s41522-017-0040-3

Drexler, K., Dannull, J., Hindennach, I., Mutschler, B., & Henning, U. (1991). Single mutations in a gene for a tail fiber component of an Escherichia coli phage can cause an extension from a protein to a carbohydrate as a receptor. *J Mol Biol, 219*(4), 655-663. doi:10.1016/0022-2836(91)90662-p

Dubberke, E. R., & Olsen, M. A. (2012). Burden of Clostridium difficile on the Healthcare System. *Clinical Infectious Diseases, 55*(suppl_2), S88-S92. doi:10.1093/cid/cis335

Duerkop, B. A., Clements, C. V., Rollins, D., Rodrigues, J. L., & Hooper, L. V. (2012). A composite bacteriophage alters colonization by an intestinal commensal bacterium. *Proc Natl Acad Sci U S A, 109*(43), 17621-17626. doi:10.1073/pnas.1206136109

Dunn, J. J., Elzinga, M., Mark, K. K., & Studier, F. W. (1981). Amino acid sequence of the gene 0.3 protein of bacteriophage T7 and nucleotide sequence of its mRNA. *Journal of Biological Chemistry, 256*(5), 2579-2585. doi:https://doi.org/10.1016/S0021-9258(19)69822-4

Dutilh, B. E., Cassman, N., McNair, K., Sanchez, S. E., Silva, G. G. Z., Boling, L., . . . Edwards, R. A. (2014). A highly abundant bacteriophage discovered in the

unknown sequences of human faecal metagenomes. *Nature communications, 5*(1), 4498. doi:10.1038/ncomms5498

Dzierzewicz, Z., Cwalina, B., Jaworska-Kik, M., Weglarz, L., & Wilczok, T. (2001). Susceptibility to antibiotics and biochemical properties of Desulfovibrio desulfuricans strains. *Acta Pol Pharm, 58*(6), 439-445.

Edwards, U., Rogall, T., Blöcker, H., Emde, M., & Böttger, E. C. (1989). Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res, 17*(19), 7843-7853. doi:10.1093/nar/17.19.7843

Ewing, W. H., Tatum, H. W., & Davis, B. R. (1958). Escherichia coli serotypes associated with edema disease of swine. *The Cornell veterinarian, 48 2*, 201-206.

Eydal, H. S., Jagevall, S., Hermansson, M., & Pedersen, K. (2009). Bacteriophage lytic to Desulfovibrio aespoeensis isolated from deep groundwater. *ISME J, 3*(10), 1139-1147. doi:10.1038/ismej.2009.66

Fairbrother, J. M., & Nadeau, É. (2019). Colibacillosis. In *Diseases of Swine* (pp. 807-834).

Fairbrother, J. M., Nadeau, É., & Gyles, C. L. (2007). Escherichia coli in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. *Animal Health Research Reviews, 6*(01), 17-39. doi:10.1079/ahr2005105

Faith, J. J., Guruge, J. L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A. L., . . . Gordon, J. I. (2013). The Long-Term Stability of the Human Gut Microbiota. *Science, 341*(6141), 1237439. doi:10.1126/science.1237439

Feng, X., Oropeza, R., Walthers, D., & Kenney, L. J. (2003). OmpR Phosphorylation and Its Role in Signaling and Pathogenesis. *ASM News, Volume 69*(8), 390-395.

Fernandes, M. A., Verstraete, S. G., Phan, T. G., Deng, X., Stekol, E., LaMere, B., . . . Delwart, E. (2019). Enteric Virome and Bacterial Microbiota in Children With Ulcerative Colitis and Crohn Disease. *J Pediatr Gastroenterol Nutr, 68*(1), 30-36. doi:10.1097/mpg.0000000000002140

Fiehn, O., Wohlgemuth, G., Scholz, M., Kind, T., Lee, D. Y., Lu, Y., . . . Nikolau, B. (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J, 53*(4), 691-704. doi:10.1111/j.1365-313X.2007.03387.x

Field, M. (2003). Intestinal ion transport and the pathophysiology of diarrhea. *Journal of Clinical Investigation, 111*(7), 931-943. doi:10.1172/jci200318326

Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., . . . Mitchell, A. L. (2017). InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Research, 45*(Database issue), D190-D199. doi:10.1093/nar/gkw1107

Fite, A. (2004a). Identification and quantitation of mucosal and faecal desulfovibrios using real time polymerase chain reaction. *Gut, 53*(4), 523-529. doi:10.1136/gut.2003.031245

Fite, A. (2004b). Identification and quantitation of mucosal and faecal desulfovibrios using real time polymerase chain reaction. *Gut, 53*(4), 523-529. doi:10.1136/gut.2003.031245

Fletcher, J. R., Pike, C. M., Parsons, R. J., Rivera, A. J., Foley, M. H., McLaren, M. R., . . . Theriot, C. M. (2021). Clostridioides difficile exploits toxin-mediated inflammation to alter the host nutritional landscape and exclude competitors from the gut microbiota. *Nature communications, 12*(1), 462. doi:10.1038/s41467-020-20746-4

Forslund, K., Hildebrand, F., Nielsen, T., Falony, G., Le Chatelier, E., Sunagawa, S., . . . Pedersen, O. (2015). Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature, 528*(7581), 262-266. doi:10.1038/nature15766

Fossum, C., Wattrang, E., Fuxler, L., Thorleif Jensen, K., & Wallgren, P. (1998). Evaluation of various cytokines (IL-6, IFN-α, IFN-γ, TNF-α) as markers for acute bacterial infection in swine - a possible role for serum interleukin-6. *Veterinary Immunology and Immunopathology, 64*(2), 161-172. doi:https://doi.org/10.1016/S0165-2427(98)00126-3

Fothergill, J. C., & Guest, J. R. (1977). Catabolism of L-lysine by Pseudomonas aeruginosa. *J Gen Microbiol, 99*(1), 139-155. doi:10.1099/00221287-99-1-139

Francis, D. H. (2002). Enterotoxigenic *Escherichia coli* infection in pigs and its diagnosis. *Journal of Swine Health Production, 10*(4), 171-175.

Franzosa, E. A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H. J., Reinker, S., . . . Xavier, R. J. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol, 4*(2), 293-305. doi:10.1038/s41564-018-0306-4

Fruciano, D. E., & Bourne, S. (2007). Phage as an antimicrobial agent: d'Herelle's heretical theories and their role in the decline of phage prophylaxis in the West. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale, 18*(1), 19-26. doi:10.1155/2007/976850

Gabay, C. (2006). Interleukin-6 and chronic inflammation. *Arthritis research & therapy, 8 Suppl 2*(Suppl 2), S3-S3. doi:10.1186/ar1917

Galkiewicz, J. P., & Kellogg, C. A. (2008). Cross-kingdom amplification using bacteria-specific primers: complications for studies of coral microbial ecology. *Appl Environ Microbiol, 74*(24), 7828-7831. doi:10.1128/AEM.01303-08

Galkin, V. E., Yu, X., Bielnicki, J., Ndjonka, D., Bell, C. E., & Egelman, E. H. (2009). Cleavage of Bacteriophage λ cI Repressor Involves the RecA C-Terminal Domain. *Journal of Molecular Biology, 385*(3), 779-787. doi:https://doi.org/10.1016/j.jmb.2008.10.081

Galpin, L., Manary, M. J., Fleming, K., Ou, C.-N., Ashorn, P., & Shulman, R. J. (2005). Effect of Lactobacillus GG on intestinal integrity in Malawian children at risk of tropical enteropathy. *The American Journal of Clinical Nutrition, 82*(5), 1040-1045. doi:10.1093/ajcn/82.5.1040

Galtier, M., De Sordi, L., Maura, D., Arachchi, H., Volant, S., Dillies, M.-A., & Debarbieux, L. (2016). Bacteriophages to reduce gut carriage of antibiotic resistant uropathogens with low impact on microbiota composition. *Environmental Microbiology, 18*(7), 2237-2245. doi:https://doi.org/10.1111/1462-2920.13284

Gao, Y., Han, F., Huang, X., Rong, Y., Yi, H., & Wang, Y. (2013). Changes in gut microbial populations, intestinal morphology, expression of tight junction proteins, and cytokine production between two pig breeds after challenge with Escherichia coli K88: A comparative study1. *Journal of Animal Science, 91*(12), 5614-5625. doi:10.2527/jas.2013-6528

García, L. R., & Molineux, I. J. (1996). Transcription-independent DNA translocation of bacteriophage T7 DNA into Escherichia coli. *Journal of Bacteriology, 178*(23), 6921-6929. doi:10.1128/jb.178.23.6921-6929.1996

Garmaeva, S., Gulyaeva, A., Sinha, T., Shkoporov, A. N., Clooney, A. G., Stockdale, S. R., . . . Zhernakova, A. (2021). Stability of the human gut virome and effect of gluten-free diet. *Cell Rep, 35*(7), 109132. doi:10.1016/j.celrep.2021.109132

Garneau, J. R., Depardieu, F., Fortier, L.-C., Bikard, D., & Monot, M. (2017). PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Scientific Reports, 7*(1), 8292. doi:10.1038/s41598-017-07910-5

Gaspari, M., Larsson, N.-G., & Gustafsson, C. M. (2004). The transcription machinery in mammalian mitochondria. *Biochimica et Biophysica Acta (BBA) -*

Bioenergetics, 1659(2), 148-152. doi:https://doi.org/10.1016/j.bbabio.2004.10.003

George, C. M., Burrowes, V., Perin, J., Oldja, L., Biswas, S., Sack, D., . . . Stine, O. C. (2018). Enteric Infections in Young Children are Associated with Environmental Enteropathy and Impaired Growth. *Trop Med Int Health, 23*(1), 26-33. doi:10.1111/tmi.13002

George, C. M., Oldja, L., Biswas, S., Perin, J., Lee, G. O., Kosek, M., . . . Faruque, A. G. (2015). Geophagy is associated with environmental enteropathy and stunting in children in rural Bangladesh. *The American journal of tropical medicine and hygiene, 92*(6), 1117-1124. doi:10.4269/ajtmh.14-0672

George, C. M., Oldja, L., Biswas, S. K., Perin, J., Lee, G. O., Ahmed, S., . . . Faruque, A. G. (2015). Fecal Markers of Environmental Enteropathy are Associated with Animal Exposure and Caregiver Hygiene in Bangladesh. *The American journal of tropical medicine and hygiene, 93*(2), 269-275. doi:10.4269/ajtmh.14-0694

Gibson, G. R., Cummings, J. H., & Macfarlane, G. T. (1991a). Growth and activities of sulphate-reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. *FEMS Microbiology Ecology, 9*(2), 103-111. doi:doi:10.1111/j.1574-6941.1991.tb01742.x

Gibson, G. R., Cummings, J. H., & Macfarlane, G. T. (1991b). Growth and activities of sulphate-reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. *FEMS Microbiology Ecology, 9*(2), 103-111. doi:doi:10.1111/j.1574-6941.1991.tb01742.x

Gill, J. J., & Hyman, P. (2010a). Phage choice, isolation, and preparation for phage therapy. *Curr Pharm Biotechnol, 11*(1), 2-14. doi:10.2174/138920110790725311

Gill, J. J., & Hyman, P. (2010b). Phage choice, isolation, and preparation for phage therapy. *Curr Pharm Biotechnol, 11*(1), 2-14. doi:10.2174/138920110790725311

Gogokhia, L., Buhrke, K., Bell, R., Hoffman, B., Brown, D. G., Hanke-Gogokhia, C., . . . Round, J. L. (2019). Expansion of Bacteriophages Is Linked to Aggravated Intestinal Inflammation and Colitis. *Cell Host Microbe, 25*(2), 285-299.e288. doi:10.1016/j.chom.2019.01.008

Goldansaz, S. A., Guo, A. C., Sajed, T., Steele, M. A., Plastow, G. S., & Wishart, D. S. (2017). Livestock metabolomics and the livestock metabolome: A systematic review. *PLoS One, 12*(5), e0177675. doi:10.1371/journal.pone.0177675

Goldstein, E. J. C., Citron, D. M., Peraino, V. A., & Cross, S. A. (2003). Desulfovibrio desulfuricans Bacteremia and Review of Human Desulfovibrio Infections. *Journal of Clinical Microbiology, 41*(6), 2752-2754. doi:10.1128/jcm.41.6.2752-2754.2003

Gomez de Agüero, M., Ganal-Vonarburg, S. C., Fuhrer, T., Rupp, S., Uchimura, Y., Li, H., . . . Macpherson, A. J. (2016). The maternal microbiota drives early postnatal innate immune development. *Science, 351*(6279), 1296. doi:10.1126/science.aad2571

Gong, W., Jia, J., Zhang, B., Mi, S., Zhang, L., Xie, X., . . . Tu, C. (2017). Serum Metabolomic Profiling of Piglets Infected with Virulent Classical Swine Fever Virus. *Frontiers in Microbiology, 8*(731). doi:10.3389/fmicb.2017.00731

Gopalakrishnan, V., Spencer, C. N., Nezi, L., Reuben, A., Andrews, M. C., Karpinets, T. V., . . . Wargo, J. A. (2018). Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science, 359*(6371), 97-103. doi:10.1126/science.aan4236

Green, S. I., Kaelber, J. T., Ma, L., Trautner, B. W., Ramig, R. F., & Maresso, A. W. (2017). Bacteriophages from ExPEC Reservoirs Kill Pandemic Multidrug-Resistant Strains of Clonal Group ST131 in Animal Models of Bacteremia. *Scientific Reports, 7*(1). doi:10.1038/srep46151

Gregory, A. C., Zablocki, O., Zayed, A. A., Howell, A., Bolduc, B., & Sullivan, M. B. (2020). The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut. *Cell Host & Microbe, 28*(5), 724-740.e728. doi:10.1016/j.chom.2020.08.003

Grimm, M. C., Elsbury, S. K., Pavli, P., & Doe, W. F. (1996). Interleukin 8: cells of origin in inflammatory bowel disease. *Gut, 38*(1), 90-98. doi:10.1136/gut.38.1.90

Grissa, I., Vergnaud, G., & Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res, 35*(Web Server issue), W52-57. doi:10.1093/nar/gkm360

Grove, A. (2013). MarR family transcription factors. *Curr Biol, 23*(4), R142-143. doi:10.1016/j.cub.2013.01.013

Guerin, E., Shkoporov, A., Stockdale, S. R., Clooney, A. G., Ryan, F. J., Sutton, T. D. S., . . . Hill, C. (2018). Biology and Taxonomy of crAss-like Bacteriophages, the

Most Abundant Virus in the Human Gut. *Cell Host & Microbe, 24*(5), 653-664.e656. doi:https://doi.org/10.1016/j.chom.2018.10.002

Guerrant, R. L., DeBoer, M. D., Moore, S. R., Scharf, R. J., & Lima, A. A. M. (2013). The impoverished gut—a triple burden of diarrhoea, stunting and chronic disease. *Nature reviews. Gastroenterology & hepatology, 10*(4), 220-229. doi:10.1038/nrgastro.2012.239

Guinane, C. M., & Cotter, P. D. (2013). Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therap Adv Gastroenterol, 6*(4), 295-308. doi:10.1177/1756283X13482996

Hagiya, H., Kimura, K., Nishi, I., Yamamoto, N., Yoshida, H., Akeda, Y., & Tomono, K. (2018). Desulfovibrio desulfuricans bacteremia: A case report and literature review. *Anaerobe, 49*, 112-115. doi:https://doi.org/10.1016/j.anaerobe.2017.12.013

Hall, A. B., Yassour, M., Sauk, J., Garner, A., Jiang, X., Arthur, T., . . . Huttenhower, C. (2017). A novel Ruminococcus gnavus clade enriched in inflammatory bowel disease patients. *Genome Med, 9*(1), 103. doi:10.1186/s13073-017-0490-5

Hanauer, S. B. (2006). Inflammatory bowel disease: epidemiology, pathogenesis, and therapeutic opportunities. *Inflamm Bowel Dis, 12 Suppl 1*, S3-9.

Handley, J., Adams, V., & Akagi, J. M. (1973). Morphology of bacteriophage-like particles from Desulfovibrio vulgaris. *J Bacteriol, 115*(3), 1205-1207.

Hannigan, G. D., Duhaime, M. B., Ruffin, M. T., Koumpouras, C. C., Schloss, P. D., & Miller, J. F. (2018). Diagnostic Potential and Interactive Dynamics of the

Colorectal Cancer Virome. *mBio, 9*(6), e02248-02218. doi:doi:10.1128/mBio.02248-18

Hargreaves, K. R., & Clokie, M. R. J. (2014). Clostridium difficile phages: still difficult? *Frontiers in Microbiology, 5*(184). doi:10.3389/fmicb.2014.00184

He, Y., Liu, Y., & Ji, P. (2021). Metabolomic Profile of Weaned Pigs Challenged with E. coli and Supplemented with Carbadox or Bacillus subtilis. *Metabolites, 11*(2). doi:10.3390/metabo11020081

Herath, D., Jayasundara, D., Ackland, D., Saeed, I., Tang, S.-L., & Halgamuge, S. (2017). Assessing Species Diversity Using Metavirome Data: Methods and Challenges. *Computational and structural biotechnology journal, 15*, 447-455. doi:10.1016/j.csbj.2017.09.001

Homolak, J., Nikolić, M., Potoč, D., Živković, M., Bakula, D., Budimir, I., . . . Vražić, D. (2021). The onset of ulcerative colitis upon Helicobacter pylori eradication in a 72-year-old woman: report of a rare case with a 3-year follow-up. *BMC gastroenterology*, *21*(1), 303-303. doi:10.1186/s12876-021-01876-5

Hsu, B. B., Gibson, T. E., Yeliseyev, V., Liu, Q., Lyon, L., Bry, L., . . . Gerber, G. K. (2019). Dynamic Modulation of the Gut Microbiota and Metabolome by Bacteriophages in a Mouse Model. *Cell Host & Microbe*, *25*(6), 803-814.e805. doi:https://doi.org/10.1016/j.chom.2019.05.001

Hu, B., Margolin, W., Molineux, I. J., & Liu, J. (2013). The Bacteriophage T7 Virion Undergoes Extensive Structural Remodeling During Infection. *Science, 339*(6119), 576. doi:10.1126/science.1231887

Hünnefeld, M., Persicke, M., Kalinowski, J., & Frunzke, J. (2019). The MarR-Type Regulator MalR Is Involved in Stress-Responsive Cell Envelope Remodeling in Corynebacterium glutamicum. *Frontiers in Microbiology, 10*(1039). doi:10.3389/fmicb.2019.01039

Huse, S. M., Ye, Y., Zhou, Y., & Fodor, A. A. (2012). A Core Human Microbiome as Viewed through 16S rRNA Sequence Clusters. *PLoS One, 7*(6), e34242. doi:10.1371/journal.pone.0034242

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J. H., Chinwalla, A. T., . . . The Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. *Nature, 486*(7402), 207-214. doi:10.1038/nature11234

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics, 11*(1), 119. doi:10.1186/1471-2105-11-119

Iida, N., Dzutsev, A., Stewart, C. A., Smith, L., Bouladoux, N., Weingarten, R. A., . . . Goldszmid, R. S. (2013). Commensal Bacteria Control Cancer Response to Therapy by Modulating the Tumor Microenvironment. *Science, 342*(6161), 967-970. doi:doi:10.1126/science.1240527

Iida, S., Streiff, M. B., Bickle, T. A., & Arber, W. (1987). Two DNA antirestriction systems of bacteriophage P1, darA, and darB: characterization of darA− phages. *Virology, 157*(1), 156-166. doi:https://doi.org/10.1016/0042-6822(87)90324-2

Iqbal, N. T., Syed, S., Kabir, F., Jamil, Z., Akhund, T., Qureshi, S., . . . Ali, A. (2019a). Pathobiome driven gut inflammation in Pakistani children with Environmental Enteric Dysfunction. *PLoS One, 14*(8), e0221095. doi:10.1371/journal.pone.0221095

Iqbal, N. T., Syed, S., Kabir, F., Jamil, Z., Akhund, T., Qureshi, S., . . . Ali, A. (2019b). Pathobiome driven gut inflammation in Pakistani children with Environmental Enteric Dysfunction. *PLoS One, 14*(8), e0221095. doi:10.1371/journal.pone.0221095

Isaacson, R. E. (1998). Enteric bacterial pathogens, villus atrophy and microbial growth. *Veterinary Quarterly, 20*(sup3), 68-72. doi:10.1080/01652176.1998.9694973

Islam, M. Z., Fokine, A., Mahalingam, M., Zhang, Z., Garcia-Doval, C., van Raaij, M. J., . . . Rao, V. B. (2019). Molecular anatomy of the receptor binding module of a bacteriophage long tail fiber. *PLOS Pathogens, 15*(12), e1008193. doi:10.1371/journal.ppat.1008193

Jack, B. R., Boutz, D. R., Paff, M. L., Smith, B. L., Bull, J. J., & Wilke, C. O. (2017). Reduced Protein Expression in a Virus Attenuated by Codon Deoptimization. *G3 (Bethesda, Md.), 7*(9), 2957-2968. doi:10.1534/g3.117.041020

Jakobsson, H. E., Jernberg, C., Andersson, A. F., Sjölund-Karlsson, M., Jansson, J. K., & Engstrand, L. (2010). Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLoS One, 5*(3), e9836. doi:10.1371/journal.pone.0009836

Jamalludeen, N., Johnson, R. P., Shewen, P. E., & Gyles, C. L. (2009). Evaluation of bacteriophages for prevention and treatment of diarrhea due to experimental enterotoxigenic Escherichia coli O149 infection of pigs. *Vet Microbiol, 136*(1-2), 135-141. doi:10.1016/j.vetmic.2008.10.021

Jamalludeen, N., Kropinski, A. M., Johnson, R. P., Lingohr, E., Harel, J., & Gyles, C. L. (2008). Complete genomic sequence of bacteriophage phiEcoM-GJ1, a novel phage that has myovirus morphology and a podovirus-like RNA polymerase. *Appl Environ Microbiol, 74*(2), 516-525. doi:10.1128/aem.00990-07

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., . . . Schmitt-Kopplin, P. (2009). Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS One, 4*(7), e6386. doi:10.1371/journal.pone.0006386

Jason, J. G., & Paul, H. (2010). Phage Choice, Isolation, and Preparation for Phage Therapy. *Current Pharmaceutical Biotechnology, 11*(1), 2-14. doi:http://dx.doi.org/10.2174/138920110790725311

Jiminez, J. A., Uwiera, T. C., Douglas Inglis, G., & Uwiera, R. R. (2015). Animal models to study acute and chronic intestinal inflammation in mammals. *Gut Pathog, 7*, 29. doi:10.1186/s13099-015-0076-y

JM, D., EP, Z., BW, W., AG, W., & JB, C. (2016). *Prevalence of Inflammatory Bowel Disease Among Adults Aged ≥18 Years - United States, 2015.* Retrieved from

Joensen, K. G., Tetzschner, A. M. M., Iguchi, A., Aarestrup, F. M., & Scheutz, F. (2015). Rapid and Easy In Silico Serotyping of Escherichia coli Isolates by Use

of Whole-Genome Sequencing Data. *Journal of Clinical Microbiology, 53*(8), 2410-2426. doi:10.1128/jcm.00008-15

Jones, G. W., & Rutter, J. M. (1972). Role of the K88 Antigen in the Pathogenesis of Neonatal Diarrhea Caused by Escherichia coli in Piglets. *Infection and Immunity, 6*(6), 918-927.

Jones, P. H., Roe, J. M., & Miller, B. G. (2001). Effects of stressors on immune parameters and on the faecal shedding of enterotoxigenic Escherichia coli in piglets following experimental inoculation. *Res Vet Sci, 70*(1), 9-17. doi:10.1053/rvsc.2000.0436

Kamimura, K., & Araki, M. (1989). Isolation and Characterization of a Bacteriophage Lytic for Desulfovibrio salexigens, a Salt-Requiring, Sulfate-Reducing Bacterium. *Appl Environ Microbiol, 55*(3), 645-648.

Kappelman, M. D., Rifas-Shiman, S. L., Kleinman, K., Ollendorf, D., Bousvaros, A., Grand, R. J., & Finkelstein, J. A. (2007). The prevalence and geographic distribution of Crohn's disease and ulcerative colitis in the United States. *Clin Gastroenterol Hepatol, 5*(12), 1424-1429. doi:10.1016/j.cgh.2007.07.012

Kazi, M., & Annapure, U. S. (2016). Bacteriophage biocontrol of foodborne pathogens. *Journal of food science and technology, 53*(3), 1355-1362. doi:10.1007/s13197-015-1996-8

Knezevic, P., Adriaenssens, E. M., & Ictv Report, C. (2021). ICTV Virus Taxonomy Profile: Inoviridae. *J Gen Virol, 102*(7). doi:10.1099/jgv.0.001614

Knights, D., Ward, T. L., McKinlay, C. E., Miller, H., Gonzalez, A., McDonald, D., & Knight, R. (2014). Rethinking “Enterotypes”. *Cell Host & Microbe, 16*(4), 433-437. doi:https://doi.org/10.1016/j.chom.2014.09.013

Kobayashi, T., Siegmund, B., Le Berre, C., Wei, S. C., Ferrante, M., Shen, B., . . . Hibi, T. (2020). Ulcerative colitis. *Nature Reviews Disease Primers, 6*(1), 74. doi:10.1038/s41572-020-0205-x

Koch, L., Poyot, T., Schnetterle, M., Guillier, S., Soulé, E., Nolent, F., . . . Biot, F. (2019). Transcriptomic studies and assessment of Yersinia pestis reference genes in various conditions. *Scientific Reports, 9*(1), 2501. doi:10.1038/s41598-019-39072-x

Korpela, K., Salonen, A., Virta, L. J., Kekkonen, R. A., Forslund, K., Bork, P., & de Vos, W. M. (2016). Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children. *Nature communications, 7*, 10410-10410. doi:10.1038/ncomms10410

Kosek, M., Guerrant, R. L., Kang, G., Bhutta, Z., Yori, P. P., Gratz, J., . . . Investigators, a. T. M.-E. N. (2014). Assessment of Environmental Enteropathy in the MAL-ED Cohort Study: Theoretical and Analytic Framework. *Clinical Infectious Diseases, 59*(suppl_4), S239-S247. doi:10.1093/cid/ciu457

Koskella, B., & Brockhurst, M. A. (2014). Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiology Reviews, 38*(5), 916-931. doi:10.1111/1574-6976.12072

Landsberger, M., Gandon, S., Meaden, S., Rollie, C., Chevallereau, A., Chabas, H., . . . van Houte, S. (2018). Anti-CRISPR Phages Cooperate to Overcome CRISPR-Cas Immunity. *Cell*, *174*(4), 908-916.e912. doi:10.1016/j.cell.2018.05.058

Langdon, A., Crook, N., & Dantas, G. (2016). The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Medicine*, *8*(1), 39-39. doi:10.1186/s13073-016-0294-z

Langmead, B., & Salzberg, S. L. (2012a). Fast gapped-read alignment with Bowtie 2. *Nature methods*, *9*(4), 357-359. doi:10.1038/nmeth.1923

Langmead, B., & Salzberg, S. L. (2012b). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357-359. doi:10.1038/nmeth.1923

Langmead, B., & Salzberg, S. L. (2012c). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357-359. doi:10.1038/nmeth.1923

Lauer, J. M., McDonald, C. M., Kisenge, R., Aboud, S., Fawzi, W. W., Liu, E., . . . Duggan, C. P. (2019). Markers of Systemic Inflammation and Environmental Enteric Dysfunction Are Not Reduced by Zinc or Multivitamins in Tanzanian Infants: A Randomized, Placebo-Controlled Trial. *The Journal of Pediatrics*, *210*, 34-40.e31. doi:https://doi.org/10.1016/j.jpeds.2019.02.016

Lavigne, R., Seto, D., Mahadevan, P., Ackermann, H. W., & Kropinski, A. M. (2008). Unifying classical and molecular taxonomic classification: analysis of the Podoviridae using BLASTP-based tools. *Res Microbiol*, *159*(5), 406-414. doi:10.1016/j.resmic.2008.03.005

Lawrence, J. G., Hatfull, G. F., & Hendrix, R. W. (2002). Imbroglios of Viral Taxonomy: Genetic Exchange and Failings of Phenetic Approaches. *Journal of Bacteriology*, *184*(17), 4891. doi:10.1128/JB.184.17.4891-4905.2002

Le Gall, G., Noor, S. O., Ridgway, K., Scovell, L., Jamieson, C., Johnson, I. T., . . . Narbad, A. (2011). Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. *J Proteome Res*, *10*(9), 4208-4218. doi:10.1021/pr2003598

Le, S., He, X., Tan, Y., Huang, G., Zhang, L., Lux, R., . . . Hu, F. (2013). Mapping the tail fiber as the receptor binding protein responsible for differential host specificity of Pseudomonas aeruginosa bacteriophages PaP1 and JG004. *PLoS One*, *8*(7), e68562-e68562. doi:10.1371/journal.pone.0068562

Lee, C. Y., Kim, S. J., Park, B. C., & Han, J. H. (2017). Effects of dietary supplementation of bacteriophages against enterotoxigenic Escherichia coli (ETEC) K88 on clinical symptoms of post-weaning pigs challenged with the ETEC pathogen. *J Anim Physiol Anim Nutr (Berl)*, *101*(1), 88-95. doi:10.1111/jpn.12513

Lee, H., Dhingra, Y., & Sashital, D. G. (2019). The Cas4-Cas1-Cas2 complex mediates precise prespacer processing during CRISPR adaptation. *eLife*, *8*, e44248. doi:10.7554/eLife.44248

Lefkowitz, E. J., Dempsey, D. M., Hendrickson, R. C., Orton, R. J., Siddell, S. G., & Smith, D. B. (2018). Virus taxonomy: the database of the International

Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research, 46*(D1), D708-D717. doi:10.1093/nar/gkx932

Leidy, J. (1853). *A Flora and Fauna Within Living Animals*. Smithsonian Institution.

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res*. doi:10.1093/nar/gkab301

Levine, J., Ellis, C. J., Furne, J. K., Springfield, J., & Levitt, M. D. (1998a). Fecal hydrogen sulfide production in ulcerative colitis. *The American journal of gastroenterology, 93*(1), 83-87. doi:10.1111/j.1572-0241.1998.083_c.x

Levine, J., Ellis, C. J., Furne, J. K., Springfield, J., & Levitt, M. D. (1998b). Fecal hydrogen sulfide production in ulcerative colitis. *Am J Gastroenterol, 93*(1), 83-87. doi:10.1111/j.1572-0241.1998.083_c.x

Lhermie, G., Grohn, Y. T., & Raboisson, D. (2016). Addressing Antimicrobial Resistance: An Overview of Priority Actions to Prevent Suboptimal Antimicrobial Use in Food-Animal Production. *Front Microbiol, 7*, 2114. doi:10.3389/fmicb.2016.02114

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics (Oxford, England), 27*(21), 2987-2993. doi:10.1093/bioinformatics/btr509

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and

SAMtools. *Bioinformatics (Oxford, England), 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Li, Y., Fu, X., Ma, X., Geng, S., Jiang, X., Huang, Q., . . . Han, X. (2018). Intestinal Microbiome-Metabolome Responses to Essential Oils in Piglets. *Frontiers in Microbiology, 9*(1988). doi:10.3389/fmicb.2018.01988

Liao, Y., Smyth, G. K., & Shi, W. (2013). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics, 30*(7), 923-930. doi:10.1093/bioinformatics/btt656

Lim, E. S., Zhou, Y., Zhao, G., Bauer, I. K., Droit, L., Ndao, I. M., . . . Holtz, L. R. (2015). Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat Med, 21*(10), 1228-1234. doi:10.1038/nm.3950

Lim, E. S., Zhou, Y., Zhao, G., Bauer, I. K., Droit, L., Ndao, I. M., . . . Holtz, L. R. (2015). Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nature Medicine, 21*(10), 1228-1234. doi:10.1038/nm.3950

Lin, D. M., Koskella, B., & Lin, H. C. (2017). Phage therapy: An alternative to antibiotics in the age of multi-drug resistance. *World journal of gastrointestinal pharmacology and therapeutics, 8*(3), 162-173. doi:10.4292/wjgpt.v8.i3.162

Lindenbaum, J., Gerson, C. D., & Kent, T. H. (1971). Recovery of Small-Intestinal Structure and Function After Residence in the Tropics. *Annals of Internal Medicine, 74*(2), 218-222. doi:10.7326/0003-4819-74-2-218

Liu, H., Geagea, H., Rousseau, G. M., Labrie, S. J., Tremblay, D. M., Liu, X., & Moineau, S. (2018). Characterization of the Escherichia coli Virulent Myophage ST32. *Viruses*, *10*(11). doi:10.3390/v10110616

Llor, C., & Bjerrum, L. (2014). Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. *Ther Adv Drug Saf*, *5*(6), 229-241. doi:10.1177/2042098614554919

Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., . . . Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*, *569*(7758), 655-662. doi:10.1038/s41586-019-1237-9

Lopez, I., Ruiz-Larrea, F., Cocolin, L., Orr, E., Phister, T., Marshall, M., . . . Mills, D. A. (2003). Design and evaluation of PCR primers for analysis of bacterial populations in wine by denaturing gradient gel electrophoresis. *Appl Environ Microbiol*, *69*(11), 6801-6807. doi:10.1128/AEM.69.11.6801-6807.2003

Loubinoux, J., Bronowicki, J.-P., Pereira, I. A. C., Mougenel, J.-L., & Le Faou, A. E. (2002a). Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *FEMS Microbiology Ecology*, *40*(2), 107-112. doi:10.1111/j.1574-6941.2002.tb00942.x

Loubinoux, J., Bronowicki, J.-P., Pereira, I. A. C., Mougenel, J.-L., & Le Faou, A. E. (2002b). Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *FEMS Microbiology Ecology*, *40*(2), 107-112. doi:10.1111/j.1574-6941.2002.tb00942.x

Lozniewski, A., Labia, R., Haristoy, X., & Mory, F. (2001). Antimicrobial susceptibilities of clinical Desulfovibrio isolates. *Antimicrobial agents and chemotherapy*, *45*(10), 2933-2935. doi:10.1128/AAC.45.10.2933-2935.2001

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature, 489*(7415), 220-230. doi:10.1038/nature11550

M.H., A. (1959). *Bacteriophages*. New York, NY, USA: Interscience Publishers.

Madec, F., Bridoux, N., Bounaix, S., Cariolet, R., Duval-Iflah, Y., Hampson, D. J., & Jestin, A. (2000). Experimental models of porcine post-weaning colibacillosis and their relationship to post-weaning diarrhoea and digestive disorders as encountered in the field. *Veterinary Microbiology*, *72*(3), 295-310. doi:https://doi.org/10.1016/S0378-1135(99)00202-3

Madsen Søren, M., Mills, D., Djordjevic, G., Israelsen, H., & Klaenhammer Todd, R. (2001). Analysis of the Genetic Switch and Replication Region of a P335-Type Bacteriophage with an Obligate Lytic Lifestyle onLactococcus lactis. *Applied and Environmental Microbiology, 67*(3), 1128-1139. doi:10.1128/AEM.67.3.1128-1139.2001

Maier, L., Pruteanu, M., Kuhn, M., Zeller, G., Telzerow, A., Anderson, E. E., . . . Typas, A. (2018). Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature, 555*(7698), 623-628. doi:10.1038/nature25979

Mainil, J. (2013). Escherichia coli virulence factors. *Vet Immunol Immunopathol*, *152*(1-2), 2-12. doi:10.1016/j.vetimm.2012.09.032

Manrique, P., Bolduc, B., Walk, S. T., van der Oost, J., de Vos, W. M., & Young, M. J. (2016). Healthy human gut phageome. *Proc Natl Acad Sci U S A, 113*(37), 10400-10405. doi:10.1073/pnas.1601060113

Mansour, L., El-Kalla, F., Kobtan, A., Abd-Elsalam, S., Yousef, M., Soliman, S., . . . Elhendawy, M. (2018). Helicobacter pylori may be an initiating factor in newly diagnosed ulcerative colitis patients: A pilot study. *World journal of clinical cases, 6*(13), 641-649. doi:10.12998/wjcc.v6.i13.641

Manuel, N., Rushing, L., Ravindran, A., Newkirk, H., Burrowes, B., Young, R., & Gonzalez, C. (2021). Complete Genome Sequence of Rhizobium japonicum Podophage Pasto. *Microbiology Resource Announcements, 10*(3), e01442-01420. doi:10.1128/MRA.01442-20

Martínez-García, E., Jatsenko, T., Kivisaar, M., & de Lorenzo, V. (2015). Freeing Pseudomonas putida KT2440 of its proviral load strengthens endurance to environmental stresses. *Environmental Microbiology, 17*(1), 76-90. doi:https://doi.org/10.1111/1462-2920.12492

Masters, B. S., Stohl, L. L., & Clayton, D. A. (1987). Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell, 51*(1), 89-99. doi:https://doi.org/10.1016/0092-8674(87)90013-4

Matson, V., Fessler, J., Bao, R., Chongsuwat, T., Zha, Y., Alegre, M.-L., . . . Gajewski, T. F. (2018). The commensal microbiome is associated with anti–PD-1

efficacy in metastatic melanoma patients. *Science, 359*(6371), 104-108. doi:doi:10.1126/science.aao3290

Mavrich, T. N., & Hatfull, G. F. (2017). Bacteriophage evolution differs by host, lifestyle and genome. *Nature Microbiology, 2*(9), 17112. doi:10.1038/nmicrobiol.2017.112

Mazmanian, S. K., Liu, C. H., Tzianabos, A. O., & Kasper, D. L. (2005). An Immunomodulatory Molecule of Symbiotic Bacteria Directs Maturation of the Host Immune System. *Cell, 122*(1), 107-118. doi:https://doi.org/10.1016/j.cell.2005.05.007

McCarville, J. L., Caminero, A., & Verdu, E. F. (2016). Novel perspectives on therapeutic modulation of the gut microbiota. *Therap Adv Gastroenterol, 9*(4), 580-593. doi:10.1177/1756283X16637819

McCracken, B. A., Spurlock, M. E., Roos, M. A., Zuckermann, F. A., & Gaskins, H. R. (1999). Weaning anorexia may contribute to local inflammation in the piglet small intestine. *J Nutr, 129*(3), 613-619. doi:10.1093/jn/129.3.613

McLamb, B. L., Gibson, A. J., Overman, E. L., Stahl, C., & Moeser, A. J. (2013). Early weaning stress in pigs impairs innate mucosal immune responses to enterotoxigenic E. coli challenge and exacerbates intestinal injury and clinical disease. *PLoS One, 8*(4), e59838-e59838. doi:10.1371/journal.pone.0059838

Mehta, R. S., Abu-Ali, G. S., Drew, D. A., Lloyd-Price, J., Subramanian, A., Lochhead, P., . . . Chan, A. T. (2018). Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol, 3*(3), 347-355. doi:10.1038/s41564-017-0096-0

Merchant, N., Lyons, E., Goff, S., Vaughn, M., Ware, D., Micklos, D., & Antin, P. (2016). The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLOS Biology, 14*(1), e1002342. doi:10.1371/journal.pbio.1002342

Mills, M., & Payne, S. M. (1995). Genetics and regulation of heme iron transport in Shigella dysenteriae and detection of an analogous system in Escherichia coli O157:H7. *Journal of Bacteriology, 177*(11), 3004-3009.

Minot, S., Bryson, A., Chehoud, C., Wu, G. D., Lewis, J. D., & Bushman, F. D. (2013). Rapid evolution of the human gut virome. *Proc Natl Acad Sci U S A, 110*(30), 12450-12455. doi:10.1073/pnas.1300833110

Minot, S., Sinha, R., Chen, J., Li, H., Keilbaugh, S. A., Wu, G. D., . . . Bushman, F. D. (2011). The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res, 21*(10), 1616-1625. doi:10.1101/gr.122705.111

Mirsepasi-Lauridsen, H. C., Du, Z., Struve, C., Charbon, G., Karczewski, J., Krogfelt, K. A., . . . Wells, J. M. (2016). Secretion of Alpha-Hemolysin by Escherichia coli Disrupts Tight Junctions in Ulcerative Colitis Patients. *Clin Transl Gastroenterol, 7*, e149. doi:10.1038/ctg.2016.3

Mollineux, I. (2006). The T7 group. In R. Calendar (Ed.), *The Bacteriophages* (2nd Edition ed., pp. 277 - 301). Oxford, United Kingdom: Oxford University Press.

Montagne, L., Boudry, G., Favier, C., Le Huerou-Luron, I., Lalles, J. P., & Seve, B. (2007). Main intestinal markers associated with the changes in gut architecture

and function in piglets after weaning. *Br J Nutr, 97*(1), 45-57. doi:10.1017/S000711450720580X

Moreno-Gallego, J. L., Chou, S.-P., Di Rienzi, S. C., Goodrich, J. K., Spector, T. D., Bell, J. T., . . . Ley, R. E. (2019). Virome Diversity Correlates with Intestinal Microbiome Diversity in Adult Monozygotic Twins. *Cell Host & Microbe, 25*(2), 261-272.e265. doi:https://doi.org/10.1016/j.chom.2019.01.019

Moxley, R. A. (2000). Edema Disease. *Veterinary Clinics of North America: Food Animal Practice, 16*(1), 175-185. doi:10.1016/s0749-0720(15)30142-0

Moxley, R. A., Berberov, E. M., Francis, D. H., Xing, J., Moayeri, M., Welch, R. A., . . . Barletta, R. G. (1998). Pathogenicity of an EnterotoxigenicEscherichia coli Hemolysin (hlyA) Mutant in Gnotobiotic Piglets. *Infection and Immunity, 66*(10), 5031-5035.

Mu, A., McDonald, D., Jarmusch, A. K., Martino, C., Brennan, C., Bryant, M., . . . Aslam, S. (2021). Assessment of the microbiome during bacteriophage therapy in combination with systemic antibiotics to treat a case of staphylococcal device infection. *Microbiome, 9*(1), 92. doi:10.1186/s40168-021-01026-9

Mutalik, V. K., Adler, B. A., Rishi, H. S., Piya, D., Zhong, C., Koskella, B., . . . Arkin, A. P. (2020). High-throughput mapping of the phage resistance landscape in E. coli. *PLOS Biology, 18*(10), e3000877. doi:10.1371/journal.pbio.3000877

Nakao, K.-i., Tanaka, K., Ichiishi, S., Mikamo, H., Shibata, T., & Watanabe, K. (2009). Susceptibilities of 23 Desulfovibrio Isolates from

Humans. *Antimicrobial agents and chemotherapy, 53*(12), 5308. doi:10.1128/AAC.00630-09

Nakatsu, G., Zhou, H., Wu, W. K. K., Wong, S. H., Coker, O. O., Dai, Z., . . . Yu, J. (2018). Alterations in Enteric Virome Are Associated With Colorectal Cancer and Survival Outcomes. *Gastroenterology, 155*(2), 529-541.e525. doi:10.1053/j.gastro.2018.04.018

Namgaladze, D., & Brune, B. (2016). Macrophage fatty acid oxidation and its roles in macrophage polarization and fatty acid-induced inflammation. *Biochim Biophys Acta, 1861*(11), 1796-1807. doi:10.1016/j.bbalip.2016.09.002

Nayfach, S., Páez-Espino, D., Call, L., Low, S. J., Sberro, H., Ivanova, N. N., . . . Kyrpides, N. C. (2021). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nat Microbiol, 6*(7), 960-970. doi:10.1038/s41564-021-00928-6

Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S., & Kyrpides, N. C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature, 568*(7753), 505-510. doi:10.1038/s41586-019-1058-x

Naylor, C., Lu, M., Haque, R., Mondal, D., Buonomo, E., Nayak, U., . . . Petri, W. A. (2015). Environmental Enteropathy, Oral Vaccine Failure and Growth Faltering in Infants in Bangladesh. *EBioMedicine, 2*(11), 1759-1766. doi:https://doi.org/10.1016/j.ebiom.2015.09.036

Nguyen, H. M., & Kang, C. (2014). Lysis delay and burst shrinkage of coliphage T7 by deletion of terminator Tφ reversed by deletion of early genes. *Journal of virology, 88*(4), 2107-2115. doi:10.1128/JVI.03274-13

Norman, J. M., Handley, S. A., Baldridge, M. T., Droit, L., Liu, C. Y., Keller, B. C., . . . Fleshner, P. (2015). Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell, 160*(3), 447-460.

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Research, 27*(5), 824-834. doi:10.1101/gr.213959.116

Nyachoti, C. M., Kiarie, E., Bhandari, S. K., Zhang, G., & Krause, D. O. (2012). Weaned pig responses to Escherichia coli K88 oral challenge when receiving a lysozyme supplement. *J Anim Sci, 90*(1), 252-260. doi:10.2527/jas.2010-3596

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J. Z., . . . Osawa, R. (2016). Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiol, 16*, 90. doi:10.1186/s12866-016-0708-5

Olsen, N. S., Forero-Junco, L., Kot, W., & Hansen, L. H. (2020). Exploring the Remarkable Diversity of Culturable Escherichia coli Phages in the Danish Wastewater Environment. *Viruses, 12*(9), 986. doi:10.3390/v12090986

Oppenheim, A. B., Kobiler, O., Stavans, J., Court, D. L., & Adhya, S. (2005). Switches in bacteriophage lambda development. *Annu Rev Genet, 39*, 409-429. doi:10.1146/annurev.genet.39.073003.113656

Osek, J. (2001). Multiplex polymerase chain reaction assay for identification of enterotoxigenic Escherichia coli strains. *J Vet Diagn Invest., 13*, 308-311.

Ott, S. J., Waetzig, G. H., Rehman, A., Moltzau-Anderson, J., Bharti, R., Grasis, J. A., . . . Schreiber, S. (2017). Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients With Clostridium difficile Infection. *Gastroenterology, 152*(4), 799-811.e797. doi:https://doi.org/10.1053/j.gastro.2016.11.010

Owen, S. V., Canals, R., Wenner, N., Hammarlöf, D. L., Kröger, C., & Hinton, J. C. D. (2020). A window into lysogeny: revealing temperate phage biology with transcriptomics. *Microbial genomics, 6*(2), e000330. doi:10.1099/mgen.0.000330

Owino, V., Ahmed, T., Freemark, M., Kelly, P., Loy, A., Manary, M., & Loechl, C. (2016). Environmental Enteric Dysfunction and Growth Failure/Stunting in Global Child Health. *Pediatrics, 138*(6).

Owusu-Asiedu, A., Nyachoti, C. M., & Marquardt, R. R. (2003). Response of early-weaned pigs to an enterotoxigenic Escherichia coli (K88) challenge when fed diets containing spray-dried porcine plasma or pea protein isolate plus egg yolk antibody, zinc oxide, fumaric acid, or antibiotic. *J Anim Sci, 81*(7), 1790-1798. doi:10.2527/2003.8171790x

Pajunen, M. I., Elizondo, M. R., Skurnik, M., Kieleczawa, J., & Molineux, I. J. (2002). Complete Nucleotide Sequence and Likely Recombinatorial Origin of Bacteriophage T3. *Journal of Molecular Biology, 319*(5), 1115-1132. doi:https://doi.org/10.1016/S0022-2836(02)00384-4

Pariente, N. (2019). A field is born. *Nature Milestones: Milestones in Human Microbiota Research*. Retrieved from https://www.nature.com/articles/d42859-019-00006-2

Pasolli, E., Asnicar, F., Manara, S., Zolfo, M., Karcher, N., Armanini, F., . . . Segata, N. (2019). Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell*, *176*(3), 649-662.e620. doi:10.1016/j.cell.2019.01.001

Pérez-Rueda, E., Collado-Vides, J., & Segovia, L. (2004). Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. *Computational Biology and Chemistry*, *28*(5), 341-350. doi:https://doi.org/10.1016/j.compbiolchem.2004.09.004

PETER K. FAGAN, M. A. H., KARL A. BETTELHEIM,, & DJORDJEVIC, S. P. (1999). Detection of Shiga-Like Toxin (stx1 and stx2), Intimin (eaeA), and Enterohemorrhagic Escherichia coli (EHEC) Hemolysin
(EHEC hlyA) Genes in Animal Feces
by Multiplex PCR. *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, *65*(2), 868-872.

Pitcher, M. C., Beatty, E. R., & Cummings, J. H. (2000a). The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. *Gut*, *46*(1), 64-72. doi:10.1136/gut.46.1.64

Pitcher, M. C., Beatty, E. R., & Cummings, J. H. (2000b). The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. *Gut, 46*(1), 64-72. doi:10.1136/gut.46.1.64

Pitcher, M. C., & Cummings, J. H. (1996). Hydrogen sulphide: a bacterial toxin in ulcerative colitis? *Gut, 39*(1), 1-4.

Piya, D., Vara, L., Russell, W. K., Young, R., & Gill, J. J. (2017). The multicomponent antirestriction system of phage P1 is linked to capsid morphogenesis. *Molecular Microbiology, 105*(3), 399-412. doi:https://doi.org/10.1111/mmi.13705

Postgate, J. R. (1984). *The Sulphate-Reducing Bacteria* (2 ed.). Great Britain, Cambridge: University Press.

Postgate, J. R., & Campbell, L. L. (1966). Classification of Desulfovibrio species, the nonsporulating sulfate-reducing bacteria. *Bacteriological reviews, 30*(4), 732-738.

Purcell, R. V., Visnovska, M., Biggs, P. J., Schmeier, S., & Frizelle, F. A. (2017). Distinct gut microbiome patterns associate with consensus molecular subtypes of colorectal cancer. *Sci Rep, 7*(1), 11590. doi:10.1038/s41598-017-11237-6

Qadri, F., Svennerholm, A.-M., Faruque, A. S. G., & Sack, R. B. (2005). Enterotoxigenic Escherichia coli in Developing Countries: Epidemiology, Microbiology, Clinical Features, Treatment, and Prevention. *Clinical Microbiology Reviews, 18*(3), 465-483. doi:10.1128/CMR.18.3.465-483.2005

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., . . . Meta, H. I. T. C. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, *464*(7285), 59-65. doi:10.1038/nature08821

Rakhuba, D. V., Kolomiets, E. I., Dey, E. S., & Novik, G. I. (2010). Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. *Pol J Microbiol, 59*(3), 145-155.

Ramalingam, R., Prasad, R. R., Shivapriya, R., & Dharmalingam, K. (2007). Molecular cloning and sequencing of mcrA locus and identification of McrA protein inEscherichia coli. *Journal of Biosciences, 17*, 217-232.

Ramsey, J., Rasche, H., Maughmer, C., Criscione, A., Mijalis, E., Liu, M., . . . Gill, J. J. (2020). Galaxy and Apollo as a biologist-friendly interface for high-quality cooperative phage genome annotation. *PLoS Comput Biol, 16*(11), e1008214. doi:10.1371/journal.pcbi.1008214

Rapp, B. J., & Wall, J. D. (1987). Genetic transfer in <em>Desulfovibrio desulfuricans</em&gt. *Proceedings of the National Academy of Sciences, 84*(24), 9128. doi:10.1073/pnas.84.24.9128

Relano, P., Mata, M., Bonneau, M., & Ritzenthaler, P. (1987). Molecular characterization and comparison of 38 virulent and temperate bacteriophages of Streptococcus lactis. *J Gen Microbiol, 133*(11), 3053-3063. doi:10.1099/00221287-133-11-3053

Ren, W., Yin, J., Gao, W., Chen, S., Duan, J., Liu, G., . . . Yin, Y. (2015). Metabolomics study of metabolic variations in enterotoxigenic Escherichia coli-infected piglets. *RSC Advances, 5*(73), 59550-59555. doi:10.1039/C5RA09513A

Rey, F. E., Gonzalez, M. D., Cheng, J., Wu, M., Ahern, P. P., & Gordon, J. I. (2013). Metabolic niche of a prominent sulfate-reducing human gut bacterium. *Proc Natl Acad Sci U S A, 110*(33), 13582-13587. doi:10.1073/pnas.1312524110

Reyes, A., Haynes, M., Hanson, N., Angly, F. E., Heath, A. C., Rohwer, F., & Gordon, J. I. (2010). Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature, 466*(7304), 334-338. doi:10.1038/nature09199

Reyes, A., Semenkovich, N. P., Whiteson, K., Rohwer, F., & Gordon, J. I. (2012). Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature Reviews Microbiology, 10*(9), 607-617. doi:10.1038/nrmicro2853

Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L., & Gordon, J. I. (2013). Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. *Proc Natl Acad Sci U S A, 110*(50), 20236-20241. doi:10.1073/pnas.1319470110

Reyes, A., Wu, M., McNulty, N. P., Rohwer, F. L., & Gordon, J. I. (2013). Gnotobiotic mouse model of phage–bacterial host dynamics in the human gut. *Proceedings of the National Academy of Sciences, 110*(50), 20236-20241. doi:10.1073/pnas.1319470110

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

Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G. A. D., Gasbarrini, A., & Mele, M. C. (2019). What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms, 7*(1). doi:10.3390/microorganisms7010014

Roda, G., Chien Ng, S., Kotze, P. G., Argollo, M., Panaccione, R., Spinelli, A., . . . Danese, S. (2020). Crohn's disease. *Nature Reviews Disease Primers, 6*(1), 22. doi:10.1038/s41572-020-0156-2

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pašić, L., Thingstad, T. F., Rohwer, F., & Mira, A. (2009). Explaining microbial population genomics through phage predation. *Nature Reviews Microbiology, 7*, 828. doi:10.1038/nrmicro2235

https://www.nature.com/articles/nrmicro2235#supplementary-information

Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Current Biology, 27*(21), R1177-R1192. doi:https://doi.org/10.1016/j.cub.2017.09.015

Roseth, A. G., Aadland, E., Jahnsen, J., & Raknerud, N. (1997). Assessment of disease activity in ulcerative colitis by faecal calprotectin, a novel granulocyte marker protein. *Digestion, 58*(2), 176-180. doi:10.1159/000201441

Routy, B., Chatelier, E. L., Derosa, L., Duong, C. P. M., Alou, M. T., Daillère, R., . . . Zitvogel, L. (2018). Gut microbiome influences efficacy of PD-1–based immunotherapy against epithelial tumors. *Science, 359*(6371), 91-97. doi:doi:10.1126/science.aan3706

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., & Barrell, B. (2000). Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England), 16*(10), 944-945. doi:10.1093/bioinformatics/16.10.944

Ryan, F. J., Ahern, A. M., Fitzgerald, R. S., Laserna-Mendieta, E. J., Power, E. M., Clooney, A. G., . . . Claesson, M. J. (2020). Colonic microbiota is associated with inflammation and host epigenomic alterations in inflammatory bowel disease. *Nature communications, 11*(1), 1512. doi:10.1038/s41467-020-15342-5

Saffarian, A., Mulet, C., Regnault, B., Amiot, A., Tran-Van-Nhieu, J., Ravel, J., . . . Pédron, T. (2019). Crypt- and Mucosa-Associated Core Microbiotas in Humans and Their Alteration in Colon Cancer Patients. *mBio, 10*(4), e01315-01319. doi:10.1128/mBio.01315-19

Samuel, G., & Reeves, P. (2003). Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydrate Research, 338*(23), 2503-2519. doi:https://doi.org/10.1016/j.carres.2003.07.009

Savage, D. C. (2001). Microbial biota of the human intestine: a tribute to some pioneering scientists. *Current issues in intestinal microbiology, 2 1*, 1-15.

Schäffler, H., & Breitrück, A. (2018). Clostridium difficile – From Colonization to Infection. *Frontiers in Microbiology, 9*(646). doi:10.3389/fmicb.2018.00646

Schaubeck, M., Clavel, T., Calasan, J., Lagkouvardos, I., Haange, S. B., Jehmlich, N., . . . Haller, D. (2016). Dysbiotic gut microbiota causes transmissible Crohn's

disease-like ileitis independent of failure in antimicrobial defence. *Gut, 65*(2), 225. doi:10.1136/gutjnl-2015-309333

Scheller, J., Chalaris, A., Schmidt-Arras, D., & Rose-John, S. (2011). The pro- and anti-inflammatory properties of the cytokine interleukin-6. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1813*(5), 878-888. doi:https://doi.org/10.1016/j.bbamcr.2011.01.034

Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J., Lessor, L., . . . Hamilton, T. (2017a). Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant Acinetobacter baumannii Infection. *Antimicrobial agents and chemotherapy, 61*(10), e00954-00917. doi:10.1128/AAC.00954-17

Schooley, R. T., Biswas, B., Gill, J. J., Hernandez-Morales, A., Lancaster, J., Lessor, L., . . . Hamilton, T. (2017b). Development and Use of Personalized Bacteriophage-Based Therapeutic Cocktails To Treat a Patient with a Disseminated Resistant Acinetobacter baumannii Infection. *Antimicrobial Agents and Chemotherapy, 61*(10), e00954-00917. doi:doi:10.1128/AAC.00954-17

Secor Patrick, R., Dandekar Ajai, A., & Garsin Danielle, A. More than Simple Parasites: the Sociobiology of Bacteriophages and Their Bacterial Hosts. *mBio, 11*(2), e00041-00020. doi:10.1128/mBio.00041-20

Seed, K. D. (2015). Battling Phages: How Bacteria Defend against Viral Attack. *PLOS Pathogens, 11*(6), e1004847-e1004847. doi:10.1371/journal.ppat.1004847

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology, 14*(8), e1002533. doi:10.1371/journal.pbio.1002533

Severinova, E., & Severinov, K. (2006). Localization of the Escherichia coli RNA polymerase beta' subunit residue phosphorylated by bacteriophage T7 kinase Gp0.7. *J Bacteriol, 188*(10), 3470-3476. doi:10.1128/jb.188.10.3470-3476.2006

Seyedirashti, S., Wood, C., & Akagi, J. M. (1991a). Induction and partial purification of bacteriophages from Desulfovibrio vulgaris (Hildenborough) and Desulfovibrio desulfuricans ATCC 13541. *J Gen Microbiol, 137*(7), 1545-1549. doi:10.1099/00221287-137-7-1545

Seyedirashti, S., Wood, C., & Akagi, J. M. (1991b). Induction and partial purification of bacteriophages from Desulfovibrio vulgaris (Hildenborough) and Desulfovibrio desulfuricans ATCC 13541. *J Gen Microbiol, 137*(7), 1545-1549. doi:10.1099/00221287-137-7-1545

Seyedirashti, S., Wood, C., & Akagi, J. M. (1992). Molecular characterization of two bacteriophages isolated from Desulfovibrio vulgaris NCIMB 8303 (Hildenborough). *J Gen Microbiol, 138*(7), 1393-1397. doi:10.1099/00221287-138-7-1393

Shafia, F., & Thompson, T. L. (1964). CALCIUM ION REQUIREMENT FOR PROLIFERATION OF BACTERIOPHAGE PHI MU-4. *J Bacteriol, 88*(2), 293-296. doi:10.1128/jb.88.2.293-296.1964

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res, 13*(11), 2498-2504. doi:10.1101/gr.1239303

Shao, Q., Trinh, J. T., & Zeng, L. (2019). High-resolution studies of lysis–lysogeny decision-making in bacteriophage lambda. *Journal of Biological Chemistry, 294*(10), 3343-3349. doi:10.1074/jbc.TM118.003209

Sheth, R. U., Cabral, V., Chen, S. P., & Wang, H. H. (2016). Manipulating Bacterial Communities by in situ Microbiome Engineering. *Trends Genet, 32*(4), 189-200. doi:10.1016/j.tig.2016.01.005

Shimada, T., Takada, H., Yamamoto, K., & Ishihama, A. (2015). Expanded roles of two-component response regulator OmpR in Escherichia coli: genomic SELEX search for novel regulation targets. *Genes Cells, 20*(11), 915-931. doi:10.1111/gtc.12282

Shis, D. L., & Bennett, M. R. (2014). Synthetic biology: the many facets of T7 RNA polymerase. *Molecular systems biology, 10*(7), 745-745. doi:10.15252/msb.20145492

Shitrit, D., Hackl, T., Laurenceau, R., Raho, N., Carlson, M. C. G., Sabehi, G., . . . Lindell, D. (2021). Genetic engineering of marine cyanophages reveals integration but not lysogeny in T7-like cyanophages. *Isme j*. doi:10.1038/s41396-021-01085-8

Shkoporov, A. N., Clooney, A. G., Sutton, T. D. S., Ryan, F. J., Daly, K. M., Nolan, J. A., . . . Hill, C. (2019). The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe, 26*(4), 527-541.e525. doi:10.1016/j.chom.2019.09.009

Shkoporov, A. N., & Hill, C. (2019). Bacteriophages of the human gut: the "known unknown" of the microbiome. *Cell Host & Microbe, 25*(2), 195-209.

Shkoporov, A. N., Khokhlova, E. V., Fitzgerald, C. B., Stockdale, S. R., Draper, L. A., Ross, R. P., & Hill, C. (2018a). ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects Bacteroides intestinalis. *Nature communications, 9*(1), 4781-4781. doi:10.1038/s41467-018-07225-7

Shkoporov, A. N., Khokhlova, E. V., Fitzgerald, C. B., Stockdale, S. R., Draper, L. A., Ross, R. P., & Hill, C. (2018b). ΦCrAss001 represents the most abundant bacteriophage family in the human gut and infects Bacteroides intestinalis. *Nature communications, 9*(1), 4781. doi:10.1038/s41467-018-07225-7

Shkoporov, A. N., Khokhlova, E. V., Stephens, N., Hueston, C., Seymour, S., Hryckowian, A. J., . . . Hill, C. (2021). Long-term persistence of crAss-like phage crAss001 is associated with phase variation in Bacteroides intestinalis. *BMC Biology, 19*(1), 163. doi:10.1186/s12915-021-01084-3

Silveira, C. B., & Rohwer, F. L. (2016). Piggyback-the-Winner in host-associated microbial communities. *NPJ Biofilms Microbiomes, 2*(1), 16010. doi:10.1038/npjbiofilms.2016.10

Singh, S. B., & Lin, H. C. (2015a). Hydrogen Sulfide in Physiology and Diseases of the Digestive Tract. *Microorganisms, 3*(4), 866-889. doi:10.3390/microorganisms3040866

Singh, S. B., & Lin, H. C. (2015b). Hydrogen Sulfide in Physiology and Diseases of the Digestive Tract. *Microorganisms, 3*(4), 866-889. doi:10.3390/microorganisms3040866

Sivan, A., Corrales, L., Hubert, N., Williams, J. B., Aquino-Michaels, K., Earley, Z. M., . . . Gajewski, T. F. (2015). Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy. *Science, 350*(6264), 1084-1089. doi:doi:10.1126/science.aac4255

Sobhani, I., Tap, J., Roudot-Thoraval, F., Roperch, J. P., Letulle, S., Langella, P., . . . Furet, J. P. (2011). Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS One, 6*(1), e16393. doi:10.1371/journal.pone.0016393

Sozhamannan, S., Chute, M. D., McAfee, F. D., Fouts, D. E., Akmal, A., Galloway, D. R., . . . Read, T. D. (2006). The Bacillus anthracis chromosome contains four conserved, excision-proficient, putative prophages. *BMC Microbiol, 6*, 34. doi:10.1186/1471-2180-6-34

Splichal, I., Fagerhol, M. K., Trebichavsky, I., Splichalova, A., & Schulze, J. (2005). The effect of intestinal colonization of germ-free pigs with Escherichia coli on calprotectin levels in plasma, intestinal and bronchoalveolar lavages. *Immunobiology, 209*(9), 681-687. doi:10.1016/j.imbio.2004.09.009

Stewart, C. J., Ajami, N. J., O'Brien, J. L., Hutchinson, D. S., Smith, D. P., Wong, M. C., . . . Petrosino, J. F. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature, 562*(7728), 583-588. doi:10.1038/s41586-018-0617-x

Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M. A., . . . Gordon, J. I. (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature, 510*(7505), 417-421. doi:10.1038/nature13421

Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M. A., . . . Gordon, J. I. (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature, 510*(7505), 417-421. doi:10.1038/nature13421

Suchodolski, J. S., Markel, M. E., Garcia-Mazcorro, J. F., Unterer, S., Heilmann, R. M., Dowd, S. E., . . . Toresson, L. (2012). The fecal microbiome in dogs with acute diarrhea and idiopathic inflammatory bowel disease. *PLoS One, 7*(12), e51907. doi:10.1371/journal.pone.0051907

Sud, M., Fahy, E., Cotter, D., Azam, K., Vadivelu, I., Burant, C., . . . Subramaniam, S. (2016). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Res, 44*(D1), D463-470. doi:10.1093/nar/gkv1042

Sugiharto, S., Hedemann, M. S., & Lauridsen, C. (2014). Plasma metabolomic profiles and immune responses of piglets after weaning and challenge with E. coli. *Journal of Animal Science and Biotechnology, 5*(1), 17. doi:10.1186/2049-1891-5-17

Sullivan, M. B., Coleman, M. L., Weigele, P., Rohwer, F., & Chisholm, S. W. (2005). Three Prochlorococcus Cyanophage Genomes: Signature Features and Ecological Interpretations. *PLOS Biology*, *3*(5), e144. doi:10.1371/journal.pbio.0030144

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics (Oxford, England), 27*(7), 1009-1010. doi:10.1093/bioinformatics/btr039

Summer, E., Liu, M., Summer, N. S., Gill, J., Janes, C., & Young, R. (2011). *Phage of sulfate reducing bacteria isolated from high saline environment*.

Summer, E. J. (2009a). Preparation of a phage DNA fragment library for whole genome shotgun sequencing. *Methods Mol Biol*, *502*, 27-46. doi:10.1007/978-1-60327-565-1_4

Summer, E. J. (2009b). Preparation of a phage DNA fragment library for whole genome shotgun sequencing. *Methods Mol Biol*, *502*, 27-46. doi:10.1007/978-1-60327-565-1_4

Summer, N., Summer, E., Gill, J., & Young, R. (2009). *Phage Remediation of Microbe Induced Corrosion*. Paper presented at the 17th International Corrosion Congress, Houston Texas.

Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., . . . Ostell, J. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research, 44*(14), 6614-6624. doi:10.1093/nar/gkw569

Taylor, V. L., Fitzpatrick, A. D., Islam, Z., & Maxwell, K. L. (2019). Chapter One - The Diverse Impacts of Phage Morons on Bacterial Fitness and Virulence. In M. Kielian, T. C. Mettenleiter, & M. J. Roossinck (Eds.), *Advances in virus research* (Vol. 103, pp. 1-31): Academic Press.

Tisza, M. J., & Buck, C. B. (2021). A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *Proc Natl Acad Sci U S A, 118*(23). doi:10.1073/pnas.2023202118

Torres, J., Hu, J., Seki, A., Eisele, C., Nair, N., Huang, R., . . . Peter, I. (2020). Infants born to mothers with IBD present with altered gut microbiome that transfers abnormalities of the adaptive immune system to germ-free mice. *Gut, 69*(1), 42. doi:10.1136/gutjnl-2018-317855

Toso, D. B., Javed, M. M., Czornyj, E., Gunsalus, R. P., & Zhou, Z. H. (2016). Discovery and Characterization of Iron Sulfide and Polyphosphate Bodies Coexisting in Archaeoglobus fulgidus Cells. *Archaea (Vancouver, B.C.), 2016*, 4706532-4706532. doi:10.1155/2016/4706532

Trehan, I., Shulman, R. J., Ou, C.-N., Maleta, K., & Manary, M. J. (2009). A randomized, double-blind, placebo-controlled trial of rifaximin, a nonabsorbable antibiotic, in the treatment of tropical enteropathy. *The American journal of gastroenterology*, *104*(9), 2326-2333. doi:10.1038/ajg.2009.270

Tsao, Y.-F., Taylor Véronique, L., Kala, S., Bondy-Denomy, J., Khan Alima, N., Bona, D., . . . O'Toole, G. Phage Morons Play an Important Role in Pseudomonas

aeruginosa Phenotypes. *Journal of Bacteriology, 200*(22), e00189-00118. doi:10.1128/JB.00189-18

Tsuda, A., Suda, W., Morita, H., Takanashi, K., Takagi, A., Koga, Y., & Hattori, M. (2015). Influence of Proton-Pump Inhibitors on the Luminal Microbiota in the Gastrointestinal Tract. *Clinical and translational gastroenterology, 6*(6), e89-e89. doi:10.1038/ctg.2015.20

Turner, D., Kropinski, A. M., & Adriaenssens, E. M. (2021). A Roadmap for Genome-Based Phage Taxonomy. *Viruses, 13*(3), 506. doi:10.3390/v13030506

Van Boeckel, T. P., Glennon, E. E., Chen, D., Gilbert, M., Robinson, T. P., Grenfell, B. T., . . . Laxminarayan, R. (2017). Reducing antimicrobial use in food animals. *Science, 357*(6358), 1350.

VanInsberghe, D., Elsherbini, J. A., Varian, B., Poutahidis, T., Erdman, S., & Polz, M. F. (2020). Diarrhoeal events can trigger long-term Clostridium difficile colonization with recurrent blooms. *Nat Microbiol, 5*(4), 642-650. doi:10.1038/s41564-020-0668-2

Vétizou, M., Pitt, J. M., Daillère, R., Lepage, P., Waldschmitt, N., Flament, C., . . . Zitvogel, L. (2015). Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science, 350*(6264), 1079-1084. doi:doi:10.1126/science.aad1329

Viaud, S., Saccheri, F., Mignot, G., Yamazaki, T., Daillère, R., Hannani, D., . . . Zitvogel, L. (2013). The intestinal microbiota modulates the anticancer immune

effects of cyclophosphamide. *Science, 342*(6161), 971-976. doi:10.1126/science.1240537

Voelker, R. (2019). FDA Approves Bacteriophage Trial. *Jama, 321*(7), 638. doi:10.1001/jama.2019.0510

Wahl, L. M., & Pattenden, T. (2017). Prophage Provide a Safe Haven for Adaptive Exploration in Temperate Viruses. *Genetics, 206*(1), 407-416. doi:10.1534/genetics.116.197541

Walker, C. B., Stolyar, S. S., Pinel, N., Yen, H. C., He, Z., Zhou, J., . . . Stahl, D. A. (2006). Recovery of temperate Desulfovibrio vulgaris bacteriophage using a novel host strain. *Environ Microbiol, 8*(11), 1950-1959. doi:10.1111/j.1462-2920.2006.01075.x

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Dempsey, D. M., . . . Davison, A. J. (2020a). Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). *Archives of Virology, 165*(11), 2737-2748. doi:10.1007/s00705-020-04752-x

Walker, P. J., Siddell, S. G., Lefkowitz, E. J., Mushegian, A. R., Adriaenssens, E. M., Dempsey, D. M., . . . Davison, A. J. (2020b). Changes to virus taxonomy and the Statutes ratified by the International Committee on Taxonomy of Viruses (2020). *Archives of Virology, 165*(11), 2737-2748. doi:10.1007/s00705-020-04752-x

Walter, J., Hertel, C., Tannock, G. W., Lis, C. M., Munro, K., & Hammes, W. P. (2001). Detection of Lactobacillus, Pediococcus, Leuconostoc, and Weissella species in human feces by using group-specific PCR primers and denaturing gradient gel

electrophoresis. *Appl Environ Microbiol, 67*(6), 2578-2585. doi:10.1128/AEM.67.6.2578-2585.2001

Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., . . . Zhao, L. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME Journal, 6*(2), 320-329. doi:10.1038/ismej.2011.109

Warren, Y. A., Citron, D. M., Merriam, C. V., & Goldstein, E. J. C. (2005). Biochemical differentiation and comparison of Desulfovibrio species and other phenotypically similar genera. *Journal of Clinical Microbiology, 43*(8), 4041-4045. doi:10.1128/JCM.43.8.4041-4045.2005

Washizaki, A., Yonesaki, T., & Otsuka, Y. (2016). Characterization of the interactions between Escherichia coli receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. *MicrobiologyOpen, 5*(6), 1003-1015. doi:https://doi.org/10.1002/mbo3.384

Wassenaar, T. M., & Gunzer, F. (2015). The prediction of virulence based on presence of virulence genes in E. coli may not always be accurate. *Gut Pathogens, 7*(1), 15. doi:10.1186/s13099-015-0062-4

Watanabe, K., & Petri, W. A. (2016). Environmental Enteropathy: Elusive but Significant Subclinical Abnormalities in Developing Countries. *EBioMedicine, 10*, 25-32. doi:https://doi.org/10.1016/j.ebiom.2016.07.030

Watanabe, K., & Petri, W. A., Jr. (2016). Environmental Enteropathy: Elusive but Significant Subclinical Abnormalities in Developing Countries. *EBioMedicine, 10*, 25-32. doi:10.1016/j.ebiom.2016.07.030

Wellock, I. J., Fortomaris, P. D., Houdijk, J. G. M., & Kyriazakis, I. (2008). Effects of dietary protein supply, weaning age and experimental enterotoxigenic Escherichia coli infection on newly weaned pigs: performance. *Animal, 2*(6), 825-833. doi:10.1017/S1751731108001559

Wesemann, D. R., Portuguese, A. J., Meyers, R. M., Gallagher, M. P., Cluff-Jones, K., Magee, J. M., . . . Alt, F. W. (2013). Microbial colonization influences early B-lineage development in the gut lamina propria. *Nature, 501*(7465), 112-115. doi:10.1038/nature12496

Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology, 13*(6), e1005595. doi:10.1371/journal.pcbi.1005595

Wilkinson, S. P., & Grove, A. (2006). Ligand-responsive transcriptional regulation by members of the MarR family of winged helix proteins. *Curr Issues Mol Biol, 8*(1), 51-62.

Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vazquez-Fresno, R., . . . Scalbert, A. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res, 46*(D1), D608-d617. doi:10.1093/nar/gkx1089

Wu, H., Liu, J., Chen, S., Zhao, Y., Zeng, S., Bin, P., . . . Zhu, G. (2018). Jejunal Metabolic Responses to Escherichia coli Infection in Piglets. *Frontiers in Microbiology, 9*(2465). doi:10.3389/fmicb.2018.02465

Xia, J., & Wishart, D. S. (2016). Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr Protoc Bioinformatics, 55*, 14 10 11-14 10 91. doi:10.1002/cpbi.11

Xiao, D., Wang, Y., Liu, G., He, J., Qiu, W., Hu, X., . . . Yin, Y. (2014). Effects of Chitosan on Intestinal Inflammation in Weaned Pigs Challenged by Enterotoxigenic Escherichia coli. *PLoS One, 9*(8), e104192. doi:10.1371/journal.pone.0104192

Yang, J., Pu, J., Lu, S., Bai, X., Wu, Y., Jin, D., . . . Xu, J. (2020). Species-Level Analysis of Human Gut Microbiota With Metataxonomics. *Frontiers in Microbiology, 11*(2029). doi:10.3389/fmicb.2020.02029

Yang, S.-C., Lin, C.-H., Aljuffali, I. A., & Fang, J.-Y. (2017). Current pathogenic Escherichia coli foodborne outbreak cases and therapy development. *Archives of Microbiology, 199*(6), 811-825. doi:10.1007/s00203-017-1393-y

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., . . . Gordon, J. I. (2012a). Human gut microbiome viewed across age and geography. *Nature, 486*(7402), 222-227. doi:10.1038/nature11053

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., . . . Gordon, J. I. (2012b). Human gut microbiome viewed across age and geography. *Nature, 486*(7402), 222-227. doi:10.1038/nature11053

Yin, J. (1993). Evolution of bacteriophage T7 in a growing plaque. *Journal of Bacteriology, 175*(5), 1272. doi:10.1128/jb.175.5.1272-1277.1993

Yin, Y., Fan, B., Liu, W., Ren, R., Chen, H., Bai, S., . . . Wang, X. (2017). Investigation into the stability and culturability of Chinese enterotypes. *Scientific Reports, 7*(1), 7947. doi:10.1038/s41598-017-08478-w

Young, R., & Gill, J. J. (2015a). Phage therapy redux—What is to be done? *Science, 350*(6265), 1163. doi:10.1126/science.aad6791

Young, R., & Gill, J. J. (2015b). Phage therapy redux—What is to be done? *Science, 350*(6265), 1163.

Yu, J., Feng, Q., Wong, S. H., Zhang, D., Liang, Q. Y., Qin, Y., . . . Wang, J. (2017). Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut, 66*(1), 70-78. doi:10.1136/gutjnl-2015-309800

Zhang, M., Sun, K., Wu, Y., Yang, Y., Tso, P., & Wu, Z. (2017). Interactions between Intestinal Microbiota and Host Immune Response in Inflammatory Bowel Disease. *Frontiers in Immunology, 8*(942). doi:10.3389/fimmu.2017.00942

Zhang, W., Berberov, E. M., Freeling, J., He, D., Moxley, R. A., & Francis, D. H. (2006). Significance of heat-stable and heat-labile enterotoxins in porcine colibacillosis in an additive model for pathogenicity studies. *Infect Immun, 74*(6), 3107-3114. doi:10.1128/IAI.01338-05

Zhang, W., Zhao, M., Ruesch, L., Omot, A., & Francis, D. (2007). Prevalence of virulence genes in Escherichia coli strains recently isolated from young pigs with diarrhea in the US. *Vet Microbiol, 123*(1-3), 145-152. doi:10.1016/j.vetmic.2007.02.018

Zhao, Y., Qin, F., Zhang, R., Giovannoni, S. J., Zhang, Z., Sun, J., . . . Rensing, C. (2019). Pelagiphages in the Podoviridae family integrate into host genomes. *Environmental Microbiology, 21*(6), 1989-2001. doi:https://doi.org/10.1111/1462-2920.14487

Zheng, D., Liwinski, T., & Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell Research, 30*(6), 492-506. doi:10.1038/s41422-020-0332-7

Zinkevich, V. V., & Beech, I. B. (2000a). Screening of sulfate-reducing bacteria in colonoscopy samples from healthy and colitic human gut mucosa. *FEMS Microbiol Ecol, 34*(2), 147-155.

Zinkevich, V. V., & Beech, I. B. (2000b). Screening of sulfate-reducing bacteria in colonoscopy samples from healthy and colitic human gut mucosa. *FEMS Microbiol Ecol, 34*(2), 147-155.

Zuo, T., Lu, X.-J., Zhang, Y., Cheung, C. P., Lam, S., Zhang, F., . . . Ng, S. C. (2019). Gut mucosal virome alterations in ulcerative colitis. *Gut, 68*(7), 1169-1179. doi:10.1136/gutjnl-2018-318131

Zuo, T., Wong, S. H., Lam, K., Lui, R., Cheung, K., Tang, W., . . . Wu, J. C. (2018). Bacteriophage transfer during faecal microbiota transplantation in Clostridium difficile infection is associated with treatment outcome. *Gut, 67*(4), 634-643.