Summer 8-2020

Innovative Approaches in the Discovery of Aquatic Mycobacteriophages

Janis H. Doss
Old Dominion University, jdoss003@odu.edu

INNOVATIVE APPROACHES IN THE DISCOVERY OF AQUATIC MYCOBACTERIOPHAGES

By

Janis H. Doss
B.S. May 2007, The Pennsylvania State University
A.A.S. May 2012, Coastal Carolina Community College

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

BIOMEDICAL SCIENCES

OLD DOMINION UNIVERSITY
August 2020

Approved by:

David Gauthier (Director)

Nazir Barekzi (Member)

Fred Dobbs (Member)

Wayne Hynes (Member)

ABSTRACT

INNOVATIVE APPROACHES IN THE DISCOVERY OF
AQUATIC MYCOBACTERIOPHAGES

Janis H. Doss
Old Dominion University, 2020
Director: Dr. David Gauthier

Bacteriophages (phages), viruses that infect bacteria, have many applications in medicine, agriculture, molecular biology, and other fields. As antibiotic resistance becomes an increasing problem, interest in phages has grown. The traditional techniques of phage discovery are successful for some phages, but others require modified procedures to achieve detectable host infection.

Mycobacterium is a diverse bacterial genus characterized by a unique cell wall containing mycolic acids, which aids in survival and pathogenesis. The aims of the present research were to isolate mycobacteriophages, use bioinformatics techniques to analyze mycobacterial prophages, and combine genetic analysis with multi-well plate host range studies to identify phages that may infect *M. marinum*, *M. fortuitum*, or *M. chelonae*, aquatic pathogens with importance to aquaculture, zoonotic infections, and more.

Firstly, three phage discovery procedures were adapted to a multi-well plate format using the host *M. smegmatis* mc^2155 as a model organism; the 96-well plate detection assay was at least as sensitive as the traditional plaque assay and a technique was developed for the purification of single phage types from mixed cultures in liquid medium. Secondly, phage enrichment from various environmental sources and

prophage induction using a variety of techniques did not result in confirmed phage isolation for several tested strains of *M. marinum*, *M. fortuitum*, and *M. chelonae*, demonstrating the difficulty of phage isolation for these species. Thirdly, the emergence and evolution of prophages was investigated in 49 sequenced genomes of the *M. ulcerans-M. marinum* complex (MuMC), resulting in 134 identified prophages in nine genomic insertion sites. Lastly, phages from a pre-existing phage collection were used in host range studies against 10 strains of mycobacteria, using both plaque assays and multi-well plate turbidity assays (MuMC, *M. fortuitum*, and *M. chelonae*).

As phages become more extensively used in medicine and other fields, it will become increasingly important to efficiently isolate phages that can infect the bacterium of interest and that have the appropriate qualities for each specific application. Phages are already known to be diverse in their morphology, physiology, and ecological roles, and it is now becoming clearer that the techniques to manipulate them must also be diverse.

This dissertation is dedicated to my husband, Mike, whose unfaltering support, love, and patience made this possible.

ACKNOWLEGMENTS

I would firstly like to thank the people who nurtured my love of science since childhood—my parents, brother, teachers, and friends who taught me critical thinking, encouraged my passions, and served as role models. Next, I would like to thank the people in my life who have given me unwavering support, love, and patience—my relatives, in-laws, friends, and especially my husband. I would also like to thank the faculty, staff, and students of ODU who have helped me along the way, whether it be by lending me supplies, teaching me a technique, helping me put holiday decorations on the lab door, or eating sushi together. Finally, I would like to thank my committee members, Dr. Nazir Barekzi, Dr. Fred Dobbs, Dr. Wayne Hynes, and especially my advisor Dr. David Gauthier. Dr. Gauthier gave me the freedom to pursue my curiosity and always had faith in my abilities; his support, guidance, and patience have allowed me to grow as a scientist and accomplish things I never would have thought possible.

NOMENCLATURE

7H9	Middlebrook 7H9 broth
AD	Albumin dextrose supplement
AFLP	Amplified fragment length polymorphism
ANOVA	Analysis of variance
attB	Bacterial attachment site
attL	Left attachment site
attP	Phage attachment site
attR	Right attachment site
BLAST	Basic Local Alignment Search Tool
bp	Base pair(s)
BRIG	BLAST Ring Image Generator
CB	Carbenicillin
CFU	Colony-forming unit(s)
CHX	Cyclohexamide
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
EPS	Extracellular polymeric substances
FDA	Food and Drug Administration
GDS	Great Dismal Swamp
HGT	Horizontal gene transfer
ICTV	International Committee on Taxonomy of Viruses

IPATH	Center for Innovative Phage Applications and Therapeutics
Kb	Kilobase pair(s)
LPS	Lipopolysaccharide
MMC	Mitomycin C
MuMC	*Mycobacterium marinum-Mycobacterium ulcerans* complex
NCBI	National Center for Biotechnology Information
NTM	Nontuberculous mycobacteria
OADC	Oleic albumin dextrose catalase supplement
OD	Optical density
ODU	Old Dominion University
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFU	Plaque-forming unit(s)
Phage	Bacteriophage
PHASTER	PHAge Search Tool – Enhanced Release
RBP	Receptor binding protein
RES	Residential area
RFLP	Restriction fragment length polymorphism
RGM	Rapidly growing mycobacteria
RNA	Ribonucleic acid
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA

SEA-PHAGES	Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science
SGM	Slow growing mycobacteria
SPSS	Statistical Package for the Social Sciences
SRA	Sequence read archive
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
USDA-NVSL-DBL	United States Department of Agriculture Animal Plant Health Inspection Service – National Veterinary Services Laboratory – Diagnostic Bacteriology Laboratory
UV	Ultraviolet
VIMS	Virginia Institute of Marine Sciences

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 AIMS

The initial goal of the project was the seemingly simple task of isolating a bacteriophage (phage) capable of infecting the fish pathogen *Mycobacterium marinum* and related species belonging to the *Mycobacterium ulcerans*/*Mycobacterium marinum* complex (MuMC) (see 1.9 Mycobacteria as Bacteriophage Hosts). These bacteria are deadly pathogens of fish; they can be problematic in aquaculture and can cause zoonotic infections of humans (1). Yet, the only publication describing possible phages of *M. marinum* describes the induction of prophages from the bacterium's genome that were observed to produce small, turbid plaques on closely related *M. marinum* strains (2); no published literature describes the isolation of environmental phages of *M. marinum*. It soon became apparent that phage isolation is dependent on many complex variables and that optimal conditions for effective phage discovery are not currently known for all bacterial host species. As a result, the project goals broadened; while still seeking an MuMC phage, it was also important to develop new phage isolation techniques that may aid in the discovery of currently un-culturable phages and to explore the history of phage infection in MuMC and other aquatic non-tuberculous mycobacteria through prophage phylogeny. The specific aims of the present research were to: 1) modify phage research techniques (enrichment, purification, and detection) for liquid medium in multi-well plates, 2) use bioinformatics techniques to analyze the

This introduction contains material adapted from:
Doss J, Culbertson K, Hahn D, Camacho J, Barekzi N. 2017. A Review of Phage Therapy against Bacterial Pathogens of Aquatic and Terrestrial Organisms. Viruses, 9: 50. doi:10.3390/v9030050

characteristics and phylogeny of mycobacterial prophages in MuMC, and 3) combine genetic analyses with multi-well plate host range studies to identify phages that may infect MuMC, *M. fortuitum*, or *M. chelonae*. Hopefully the findings of these studies will benefit phage discovery in a number of hosts and will encourage others to explore alternate isolation techniques to gain a wider understanding of the full diversity of phages.

1.2 HISTORY

Phages, which are obligate viral parasites of bacteria, were independently discovered by Frederick Twort in 1915 and by Félix d'Hérelle in 1917 (3, 4). Early excitement over phages led to numerous publications and products, but due to problems of reproducibility stemming from poor experimental design and a lack of understanding of the biology of phages, eventually the West turned its interest towards the more predictable chemical antibiotics (5). However, phages were still appreciated for other useful applications and were continuously integral in the advances of molecular biology, evolution, and ecology (6). In addition, interest in phage therapy has increased throughout the world due to antibiotic resistance in pathogenic bacteria (6). Now, phages are used in many diverse fields such as agriculture, medicine, biotechnology, and education (5). A summary of important discoveries involving phages is found in Table 1.

TABLE 1 Important scientific achievements involving phages

Year	Discovery
1915	Frederick Twort describes a bacteriolytic agent; likely the first observation of phage plaques (5)
1917	Félix d'Hérelle independently re-discovers phages; uses the term “bacteriophage” and pursues phage therapy (5)
1920-1930	Early phage therapy research and sales of phage treatments, but poor experimental design and lack of understanding of phage biology impedes success (5)
1939	Emory Ellis and Max Delbrück conduct the one-step growth experiment (identified key concepts in the phage lytic cycle, including latent period and burst size) (5)
1930-1940	Success of chemical antibiotics and concerns over the efficacy of phage therapy; phage therapy research stalls in the West but continues in Eastern Europe and Russia (5)
1940	Phages visualized by electron microscopy (5)
1940-1950	Beginning of the “phage group,” a collaboration of Max Delbrück, Salvador Luria, Alfred Hershey, and others that made advances in molecular biology using phages as model organisms (6)
1943	Salvador Luria and Max Delbrück: mutations arise randomly even in the absence of selective pressures (5)
1945	Salvador Luria: phages can mutate to overcome host resistance (6)
1946	Alfred Hershey and Max Delbrück: two phages can genetically recombine during co-infection (6)
1950	André Lwoff, Louis Siminovitch, and Niels Kjeldgaard: prophage induction via ultraviolet irradiation (6)
1951	Victor Freeman: phages can transfer virulence factors to their hosts (6)
1952	Alfred Hershey and Martha Chase: 'Waring blender experiment'; DNA is the hereditary material (5)
1952	Joshua Lederberg and Norton Zinder: transduction demonstrated (5)
1952	Salvador Luria and Mary Human: after passage through a host, the phage’s ability to infect other hosts decreases (led to the discovery of restriction enzymes) (6)
1955	Seymour Benzer: using the phage T4 *r*II locus, demonstrated that genes have a linear structure (5)
1961	Francis Crick, Sydney Brenner, Leslie Barnett, and R.J. Watts-Tobin: nucleotides are read as codons (6)
1976	Walter Fiers: first genome sequenced (MS2, an RNA phage) (5)
1985	George Smith: invention of phage display, a technique to study protein interactions by displaying them on the surface of filamentous phages (5)

Adapted from reviews by Salmond and Fineran (5) and Keen (6).

1.3 STRUCTURE AND TAXONOMY

Viral taxonomy is complex and sometimes controversial; the evolutionary history of viruses is often ambiguous due to horizontal gene transfer (HGT) and the lack of a universal, conserved gene (comparable to the bacterial 16S rRNA gene) that could serve as a basis for taxonomic classification (7). The International Committee on Taxonomy of Viruses (ICTV) assigns official taxonomic designations to viruses and revises their classification systems periodically (7). Although viral taxonomy was formerly based primarily on morphology, genomic comparisons are increasingly emphasized and many taxonomic designations are in the process of reorganization (7).

Approximately 96% of phages visualized by electron microscopy belong to the order *Caudovirales* (the tailed phages) (Fig. 1) (5). Most members of this order fall into three families: *Myoviridae* have contractile tails, *Siphoviridae* have flexible, non-contractile tails, and *Podoviridae* have shorter, non-contractile tails (5). Two families were recently added to the *Caudovirales* order, *Ackermannviridae* and *Herelleviridae*, which had previously been classified as *Myoviridae* (8). *Ackermannviridae* differs from *Myoviridae* in that pronged structures and short filaments with rounded tips protrude from the baseplate (8). *Herelleviridae* vary genetically but not structurally from *Myoviridae* (8).

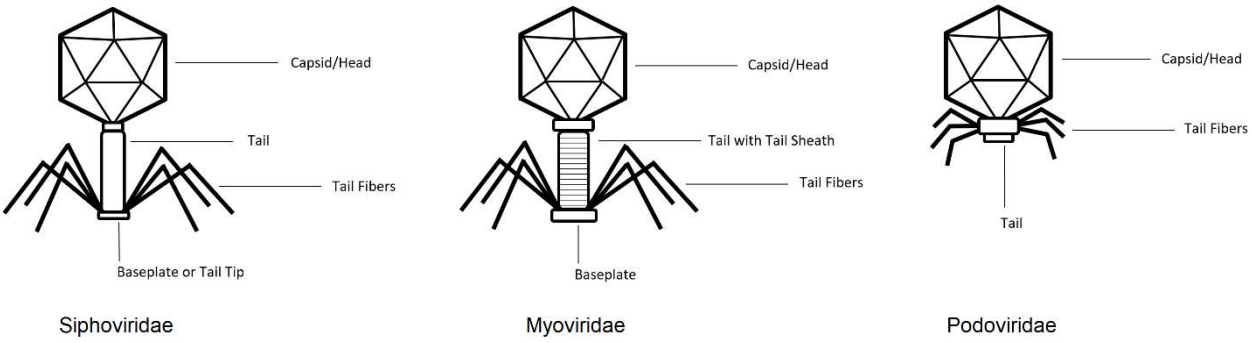

FIG 1 Typical morphology and structural characteristics of the most frequently assigned phage families (*Siphoviridae*, *Myoviridae*, and *Podoviridae* of the order Caudovirales). *Siphoviridae* phages are characterized by a long but non-contractile tail, while *Myoviridae* phages have a tail sheath around the tail tube, which allows contraction (8, 9). *Podoviridae* phages are characterized by a short, non-contractile tail (8, 9).

Phage capsids can be polyhedral (especially icosahedral) or filamentous (8). *Caudovirales* capsids are usually (about 75%) icosahedral and range in diameter from 45 nm to 185 nm (8). Capsid size correlates with genome size, since phage genomes are packed into capsids at a fairly consistent physical density (10). The major capsid protein fold, an important structural motif ensuring proper formation of the capsid, is conserved among *Caudovirales* phages, archaeal viruses, and adenoviruses, yet shares no amino acid or nucleotide sequence similarity (8). The capsid connects to the tail via a portal protein complex, which has the shape of a dodecameric ring (8). Like the capsid protein fold, *Caudovirales* portal proteins have high structural similarity despite a low sequence similarity; this pattern is also found in capsid-tail connector proteins and receptor binding proteins (RBPs) (8). The tail of a *Siphoviridae* phage is composed of a

tapemeasure protein within a tail tube (8). In *Myoviridae* phages, a sheath surrounds the central tail structure and is responsible for contracting the tail during injection (8). The tail often terminates in a baseplate and tail fibers or tail spikes (11). *Myoviridae* phages always have a baseplate, while *Siphoviridae* phages sometimes have a simple baseplate and *Siphoviridae* phages that infect Gram-negative hosts have no baseplate (9). In *Podoviridae* phages, the tail is typically composed of an upper tail adaptor protein, a tail tube, a lower nozzle, and tail fibers or tail spikes (9). Tail spikes contain enzymes for cell wall degradation; tail fibers are similar to tail spikes but are longer, more complex, and do not contain enzymes (9).

The genetic material of a phage can be single stranded RNA (ssRNA), double stranded RNA (dsRNA), single stranded DNA (ssDNA), or double stranded DNA (dsDNA) (11). *Caudovirales* phages tend to have linear, double-stranded DNA (5). Phage genome size ranges from less than 4 Kb to over 400 Kb (6). The smallest phage genomes are the 3.3 Kb genomes of ssRNA *Escherichia coli* phages (10). The smallest *Siphoviridae* genome is the 14.2 Kb genome of the *Rhodococcus* phage RRH1, which encodes only 20 genes (12). Some recently discovered phages have genome sizes up to 735 Kb and contain such a high number and diversity of genes that their genomes bear more of a resemblance to the genomes of small, obligately parasitic bacteria than to phage genomes (13). These “huge phages,” “jumbo phages,” or “mega phages” often contain clustered regularly interspaced short palindromic repeats (CRISPR)-Cas systems and genes involved in metabolism, transcription, and translation (13).

The major determinant of host specificity in phages are RBPs, which interact with bacterial membrane surface receptors (11). If a phage can recognize more than one

type of receptor, it can have a broader host range; however, phages with broader host ranges tend to have smaller burst sizes (the average number of progeny phages that are produced with each infection) (14). Host range is also limited by the differences in cell envelope composition among bacteria; for instance, no phages are currently known to infect both Gram-positive and Gram-negative bacteria (9). The surface receptor to which a phage binds can be any molecule on the cell surface, whether protein, carbohydrate, or lipid (15). Common receptors include peptidoglycan, teichoic acids, polysaccharides, and lipopolysaccharide (LPS) (15). If the receptor is a protein, the tail fibers will have a sharp or spiked end, while a carbohydrate receptor will have a larger baseplate (15). RBPs are diverse in structure, location, and mechanism: in *Myoviridae*, RBPs are typically found on the long and short tail fibers, in *Siphoviridae* they are usually found on the baseplate, and in some phages they are found on the tail spikes (15).

1.4 LIFECYCLE

Based on lifecycle strategy, phages can be separated into two categories: lytic phages can only undergo the lytic cycle, in which the host cell is lysed to release progeny phages, while temperate phages are capable of both the lytic cycle and the lysogenic cycle (16). In the lysogenic cycle, the phage integrates its genome into the host genome; the resulting DNA segment is referred to as a prophage and the cell containing it is called a lysogen (16).

The lytic cycle begins with the phage binding to host surface receptors (adsorption), penetrating the cell membrane, and injecting its genetic material (5). No

chemical energy is used in the attachment of the phage to its host or the injection of genetic material since viruses lack the biochemical machinery for energy usage; only diffusion and Brownian motion bring the phage to its host (11). Some phages carry enzymes that can degrade LPS, peptidoglycans, capsules, or biofilms (11). During adsorption, the phage first binds reversibly to a receptor, then it binds irreversibly to the same receptor or to a different receptor (15). Mechanisms of adsorption vary greatly among phages and can be simple or complex (15). Many mechanisms require the presence of calcium to mediate conformational changes (15).

The baseplate coordinates contraction of the tail fibers to aid binding (11). Distal tail proteins help form the center of the baseplate and links the baseplate to the tail (15). The upper baseplate protein links the distal tail proteins to the RBPs (15). The change in orientation of the tail fibers relative to the baseplate signals the baseplate to undergo a conformational change that induces sheath contraction in *Myoviridae phages* (11). The spikes of the baseplate penetrate the host membrane and can release enzymes that aid in injection (11). Tail-associated lysin protein degrades peptidoglycan (15). The tail tube channel opens and releases the phage's genetic material into the host cytoplasm (11). The host's replication machinery is used to produce phage components and enzymes, then progeny phages assemble (16).

Bacteria often undergo physiological changes in response to phage infection, especially during the virion assembly and enzyme production stages in late phage infection (17). Metabolomic, proteomic, and transcriptomic analyses indicate that changes occur in metabolism, biosynthesis, stress responses, and gene expression (17). Phages often redirect host resources and biochemical machinery away from host

functions in order to expedite phage production (14). Host nucleic acids may also be recycled into phage nucleic acids (14). Nucleic acid packaging is coordinated by the terminase protein (5). Holins perforate the host cell membrane, allowing lysins to degrade the peptidoglycan layer; lysis of the host then releases the progeny phages (11).

The lysogenic cycle also begins with adsorption, penetration, and injection (5). The phage genome integrates into the host chromosome (in most cases) via the action of the integrase gene, which can target a specific site in the genome or integrate randomly (16, 18). Integration occurs between the bacterial attachment site (*attB*) in the host genome and the phage attachment site (*attP*) in the phage genome (18). Repressor genes prevent premature induction and can also be responsible for superinfection exclusion (protection from future phage infection) (16).

Prophage DNA is replicated along with the host DNA whenever binary fission occurs (16). By remaining in the host genome, the prophage risks inactivation due to the accumulation of random mutations (such as insertions, deletions, point mutations, and transpositions) which can render essential phage genes inoperative (16). If the prophage contains genes beneficial to the host (termed “morons”), selective pressures may delay or prevent this genetic degradation (16). Prophages are commonly found in bacterial genomes in both active and degraded forms and can make up as much as 20% of the genome (10).

The prophage may at some point re-enter the lytic cycle (induction), either spontaneously or due to stress on the lysogen (18). Stressors, such as ultraviolet light, mitomycin C (MMC), and certain antibiotics, usually directly or indirectly lead to DNA

damage (16). Whether a temperate phage enters the lytic cycle or the lysogenic cycle upon host infection does not appear to be a random phenomenon; in phages in which this topic has been investigated, numerous complex factors can influence the choice, including transcriptional activators and repressors, cellular proteases, transcription antitermination, the metabolic state of the host cell, and even chemical signals from other phages (arbitrium) (19).

Rarely, some prophages do not integrate into the host chromosome and instead form plasmids (18). Another rare occurence, pseudolysogeny, occurs when severe host stress prevents replication or integration and the phage genome remains inactive in the host cytoplasm until it can resume the lytic or lysogenic cycle (18).

1.5 GENETICS AND EVOLUTION

Phage-bacterium interactions can influence evolution in both phages and their hosts via diverse mechanisms. Phage resistance mechanisms, such as CRISPR-Cas systems, restriction modification systems, receptor modifications, abortive infection systems, and Argonaute proteins, are countered by phage adaptations, resulting in an evolutionary arms race (18, 19). Phages can avoid the action of CRISPR-Cas systems by preventing formation of the protein complex or mutating their genetic sequence, can prevent abortive infection by producing antitoxins, and can mutate restriction sites or methylate their genetic sequence to block restriction endonucleases (20). Phages can also respond to host resistance by broadening their host range (20). Bacteria can change or remove surface molecules that act as phage receptors; however, these molecules may have important functions (such as virulence) and changing them may

reduce the fitness of the bacterium (20). Receptors can be masked with specialized proteins or a thick extracellular matrix (20). However, a bacterium cannot indefinitely accumulate anti-phage modules; gaining resistance to a phage can reduce fitness or even increase a bacterium's susceptibility to other phages (14).

Temperate phages may evolve more mutualistically with their hosts since their survival depends on at least temporary survival of the host (18). Expression of prophage genes by the host cell is called lysogenic conversion (16). Surprisingly, prophage genes are sometimes regulated by bacterial transcription factors (18). The expression of phage genes can encourage maintenance of the prophage via selective pressures, by enhancing such host functions as adhesion, invasion, biofilm development, growth rate, and antimicrobial resistance (16). Some bacterial strains are only virulent when carrying a prophage encoding a toxin, for example diptheria toxin (*Corynebacterium diptheriae*), Shiga toxin (Shigatoxigenic *E. coli*), and Ctx cytotoxin (*Vibrio cholerae*) (16). If prophage insertion disrupts an essential bacterial gene, the prophage may provide a copy of the gene or a portion that completes the gene (18). In active lysogeny, the disrupted gene is restored by prophage excision which may be reversible or irreversible; the prophage acts as a regulatory mechanism for the expression of the gene (18).

Frequent HGT events result in genomic mosaicism, making phylogenetic analysis and taxonomy difficult (10). Mosaicism produces phage genomes that have similarity to other phages in a patchwork pattern, with abrupt transitions, across their genomes (8). In other words, a hypothetical phage A might have high sequence similarity to phage B's capsid genes, phage C's tail genes, and phage D's integrase gene, but have little to no sequence similarity to those phages at any other loci. Some of

the genes in a phage genome may not be similar to genes of any known phage, for instance if the gene was obtained relatively recently from a bacterium or other organism (20). There is even a case of a phage with a series of genes believed to be of eukaryotic origin: the genome of the *Wolbachia* phage WO contains a eukaryotic association module that includes a black widow spider toxin (21). Groups of genes that encode proteins that must interact directly with each other, such as capsid genes, tail genes, and lysis genes, tend to be less mosaic within their group, in other words these genes tend to be transferred either as a group or not at all (10). In fact, these genes tend to be found spatially close in the genome (10).

Transduction is the transfer of nucleic acids into bacteria via phages and can result in the transfer of bacterial DNA from one individual to another bacterium (20). In generalized transduction, bacterial DNA is added into capsids instead of phage genetic material and can be transferred to another bacterium (20). In specialized transduction, bacterial DNA is added to capsids with the phage genetic material during prophage induction (20).

Recombination between phages can occur when more than one phage enters a host cell; this can occur between two phages in the lytic cycle, two phages in the lysogenic cycle, or one of each (8). The types of recombination that typically occur between phages are homologous recombination, non-homologous recombination, and homeologous (relaxed) recombination (8). Non-homologous recombination occurs randomly, homeologous recombination occurs between similar sequences, and homologous recombination is catalyzed by recombination machinery encoded by the phage (8).

1.6 ECOLOGY

Although the relationship between a phage and its host bacterium is usually described as a parasite-host or predator-prey interaction, depending on the phage and host it may be a neutral, commensal, or even mutualistic relationship (i.e. lysogenic conversion or superinfection exclusion) (20). Phages and bacteria can even be competitors in the case of predatory bacteria targeting a potential phage host (20).

The diversity, distribution, and abundance of phages can be ascertained from cultured phages, metagenomic data, and prophage data (10). Some phages are abundant globally while others are found in specific environments (22).

In marine environments, phage abundance and diversity vary by geographical location, water depth, distance from shore (8). Marine phages have profound impacts on biogeochemical cycling and microbial population structure (23). Soil phages are also ecologically important, and their abundance and diversity depend on factors such as biome, soil type, altitude, pH, and host abundance (8). In the human gut, *Caudovirales* and *Microviridae* predominate; phage composition varies by individual and their health status but is generally stable after early childhood (8).

1.7 APPLICATIONS

The properties of phages can be exploited in many fields, including education, biotechnology, surface disinfection, food safety, and medicine (6).

The Science Education Alliance–Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program uses phage discovery to provide

undergraduate students with an authentic research experience to increase their interest in and understanding of the process of science (24). Phage research allows the students to gain experience in a variety of methods such as bacterial culture, sterile technique, and DNA extraction, while also gaining widely applicable skills such as problem solving, persistence, and teamwork (24, 25). Out of the 12,377 phage genomes in the National Center for Biotechnology Information (NCBI) database in July 2020, 1,800 of them were *M. smegmatis* phages, most of which were isolated by the SEA-PHAGES program (26). The SEA-PHAGES program has not only been a positive experience for students, it has also led to significant contributions to the knowledge of phage diversity, genetics, and evolution (27).

In biotechnology, the many interesting properties of phages can be harnessed in unexpected ways. The structural characteristics of phages can be utilized for phage display, phage vaccines, and nanocarriers (11). Phage display is a technique usually used to develop vaccines and other therapeutic agents whereby filamentous phages are engineered to display proteins or peptides on their surface, producing a library that can be screened for affinity for the target molecule (11). Phage vaccines can be made by covering the surface of a phage in the antigen of interest through genetic engineering (11). Phage capsids filled with haemagglutinin ligands can bind to and cover another viral capsid, thus inactivating it (28). Phage capsids can be used as nanocarriers to bring drugs or gene therapy to a specific location by attaching targeted ligands to the capsid surface (11). The binding ability of phages can be used to tether cells to reporter molecules or to surfaces (11). The ability of a phage to carry out transduction makes them useful as cloning vectors (5). Other characteristics such as short generation time,

small size, and ease of manipulation make phages ideal model organisms in the study of genetics, evolution, ecology, and other subjects (5).

Phages can be used for surface disinfection, especially for those having biofilms (5). Biofilms are dynamic aggregates of microbes defined by their complex spatial structure and production of extracellular polymeric substances (EPS) that help to hold the cells together (29). When living in a biofilm, bacteria are more resistant to the stresses of dehydration, pH, salinity, and ultraviolet light (30). The resistance of bacterial biofilms to chemical antibiotics and the promotion of pathogenicity that biofilms confer are especially relevant to the field of medicine (29). Both phage infection and prophage presence have been widely studied in relation to biofilm production, maintenance, and elimination, but results have been inconsistent (17, 29). A phage's interaction with a biofilm is influenced by multiple factors such as phage characteristics (burst size, latent period, and production of EPS-dissipating enzymes), bacterial strain(s) present and their metabolic state, EPS composition, and environmental conditions (29). In agriculture and food safety, phages can be used to prevent and/or reduce bacterial contamination in plants, animals, fungi, and the products derived from them (11). Phages can be used for biodetection of pathogens in food, water, and patient samples; a fluorescent or bioluminescent reporter gene can be engineered into the phage genome for easier reading of the assay (11). Surface disinfection and biofilm reduction can also be carried out using phage-derived enzymes (11).

Phage therapy is the use of intact phage particles for the treatment of bacterial infections (31). Phage therapy was used experimentally and clinically in the early 20th century, but due to the increased reliability of chemical antibiotics the United States and

Western Europe lost interest in phage therapy until the crisis of antibiotic resistance necessitated the development of alternate treatments (32). There is not yet any phage approved for therapeutic use in humans by the United States Food and Drug Administration (FDA), but patients in the U.S. have a few options (32). Unapproved treatments can be given to a patient under compassionate use laws, but the patient must meet certain criteria (essentially, failure of conventional treatment and no available clinical trials) (32). The patient could travel to a clinic in a country in which phage therapy is approved, for instance, the Eliava Institute in Tbilisi, Georgia or the Ludwik Hirszfeld Institute of Immunology and Experimental Therapy in Poland (32). The Center for Innovative Phage Applications and Therapeutics (IPATH) at the University of California San Diego School of Medicine is the first phage therapy center in North America, providing phage preparations for compassionate use cases in the United States (32). Phages for compassionate use typically originate from academic laboratories, phage banks, or biotechnology companies (32). The Phage Directory is an organization that coordinates the crowd-sourcing of phages for compassionate use cases (32).

Phages are beneficial as a therapeutic agent because they are selective (only kill bacteria in their host range), propagate exponentially, and are generally safe (33). Phages used in phage therapy must have a strictly lytic lifestyle, propagate reliably and rapidly, and not cause an immunologic reaction (31). The host bacterium must be carefully evaluated to ensure that it does not carry toxins that could be released upon lysis or genes likely to be transmitted via the phage to other bacteria, such as virulence factors within a prophage sequence (33). The threat of resistance is reduced by using

phage cocktails of, ideally, distantly related phages that bind to different receptors, and changing phages in the cocktail as needed (either from a pre-existing collection or isolated from the environment) (33). Phage therapy preparations can be administered in many ways, depending on the type of infection: by aerosol, orally as a solid or liquid, by topical application, or by injection (11). The phage cocktail must be prepared carefully and then purified to prevent contamination, especially of endotoxins such as LPS (33).

One of the barriers to the acceptance of phage therapy is the tendency to evaluate phage preparations based on the characteristics developed for the evaluation of chemical antibiotics (11). The pharmacokinetics (amount of active substance in the system over time) and pharmacodynamics (the physiological effects of the substance over time) can be very different for phages than for traditional antibiotics (11). For example, the number of phages present in a patient at a given time cannot be easily calculated; it will be affected by many factors such as the starting dosage, sensitivity of the phage to pH and enzymes, rate of absorption, distribution of phages in the patient's body, number of susceptible bacteria over time, phage resistance mechanisms and evolution of the host bacterium, phage clearance rate, immune interaction, latent period, and burst size (11, 33). Using phage therapy concurrently with chemical antibiotics can have synergistic effects, especially in the treatment of biofilm-producing bacterial infections (33).

Phage-derived products such as lysins and depolymerases can be used in place of whole phages, retaining the specificity and bactericidal properties of phage therapy without many of the concerns associated with using a biological particle that can propagate inside the patient, enter the environment, or undergo HGT (33).

Unfortunately, phage enzymes stimulate a much greater immune response than whole phages and are rapidly cleared from the body (33).

1.8 CHALLENGES AND APPROACHES

Metagenomic studies have revealed that phages are likely much more diverse than previously believed (8). “The great plaque count anomaly” refers to the great discrepancy between phages known to exist by microscopy and metagenomics and the actual number of phages that have been cultured (34). There is no single technique that can fully define the phages present in a sample, but each has its benefits: microscopy can determine phage morphology, metagenomics can provide genetic information on a large number of phages even when host and appropriate culture conditions are unknown, and culture techniques can establish phenotypic traits such as plaque size, host range, and morphology, in addition to providing a high titer phage solution for future experimentation (22). A combination of these techniques must be used to gain a complete picture of the phage composition of a sample (22).

There are some bacterial hosts whose phages are notoriously difficult to propagate. The difficulty of isolating phages of *Clostridium difficile* has been attributed to the high rates of lysogeny of its phages, which could be selected for by superinfection exclusion due to the large number of prophages in genomes of this species (35). All currently sequenced *C. difficile*-infecting phages contain genes associated with lysogeny, and no strictly lytic phages are known to infect this species (35). Decreased rates of host lysis due to phages favoring the lysogenic cycle makes both plaque and turbidity detection assays difficult. *Bifidobacterium* phages are known to exist because

of studies using PCR and electron microscopy, and because of the presence of prophages in their genomes; however, no bifidobacterial phages have been observed to produce visible plaques (36). Even bifidobacterial prophages that seem to be fully intact and inducible have not been successfully induced in the laboratory (37), demonstrating that a history of phage infection may be inferred even if phage culture is not possible. The possibility of non-typical lifestyles may complicate matters further; although many archaeal viruses have similar lifestyles to those of phages, some appear to be released from their host cell without cell lysis (38).

There are many reasons why a phage may not be detectable by traditional methods. Agar is a greater barrier to phage diffusion than a liquid medium and some phages may require a lower agar concentration to form visible plaques (39). Other changes may need to be made to the standard agar composition for certain phages: some *Lactococcus lactis* phages produce larger plaques with the addition of glycine (40), some *E. coli* phages only produced clearly visible plaques with the addition of a low concentration of ampicillin (41), and *Flavobacterium columnare* is more susceptible to phage infection in the presence of mucins (42). Larger phages are especially likely to be under-cultured and often require lower agar concentrations and alternates to the traditional filtering step (34). Other factors that can influence plaque presence and size include host concentration, adsorption efficiency, latent period, and burst size (39). In aquatic phages, infection efficiency is influenced by salinity, nutrient content, and to a lesser extent pH and temperature (43). Difficulties may also occur if the phage is not well adapted to the bacterial species used in the assay; the preferred host of a phage

can be different from the host that was used for phage isolation and the optimal host may not be known (10).

1.9 MYCOBACTERIA AS PHAGE HOSTS

Mycobacteria are acid-fast, aerobic to microaerophilic, non-motile bacilli with a high G+C content (57 – 73%) and characteristic mycolic acids in their cell envelope (44). Almost 200 species are found in the genus *Mycobacterium*, which is in the family *Mycobacteriaceae* and the order *Actinomycetales* (44). This genus includes obligate human pathogens, such as the causative agents of tuberculosis and leprosy, pathogens of non-human animals, and saprophytes (44).

The Runyon classification organizes mycobacteria into groups based on phenotypic characteristics (45). Slow growing mycobacteria (SGM) take more than seven days to form colonies while rapidly growing mycobacteria (RGM) form colonies in less than seven days (45). The SGM and RGM are, for the most part, from different phylogenetic branches of the genus *Mycobacterium* (44). SGM usually contain one copy of the ribosomal operon and RGM usually contain two copies, although there are several exceptions (44). The SGM are further classified as photochromogenic (producing pigments in response to light), scotochromogenic (always pigmented), and non-photochromogenic (never producing pigments) (45). Phylogenetically, RGM form three main clades: one clade contains *M. fortuitum*, one clade contains *M. smegmatis*, and one clade is made up of the *M. abscessus-M. chelonae* group (44). SGM includes the *M. avium complex* and *M. tuberculosis* complex clades, in addition to many other groups (44). Genomic analyses indicate that some species could be combined based

on nucleotide identity; many mycobacterial taxonomies are currently in a state of flux while a clearer view of their lineages is achieved (46). Complicating the mycobacterial phylogeny is the sometimes unclear role of HGT in the evolution of these species (46). A recent proposition to split the *Mycobacterium* genus into five genera is controversial (44).

Several cultured phages have potential in treating mycobacterial pathogens, including *M. tuberculosis*, *M. bovis, M. avium, and M. ulcerans* (47). Phage therapy was recently used to treat a cystic fibrosis patient with an *M. abscessus* lung infection, resulting in considerable clinical improvement (48).

Tuberculosis is a difficult-to-treat respiratory disease that kills millions of people each year; an estimated 23% of people worldwide have a latent tuberculosis infection (49). Several features make *M. tuberculosis* difficult to treat: the acid-fast cell envelope, the ability to live intracellularly in macrophages, and the ability to lie dormant in granulomas (49). One option is to use *M. smegmatis* as a carrier to bring a phage to *M. tuberculosis* (49). Another option is using phage-derived enzymes such as lysins, which could be fused to an engineered protein that delivers the lysins into human cells (49). Phages can also be used in the diagnosis of tuberculosis by performing a phage amplification assay or using the phage as a carrier for a reporter gene (50).

Nontuberculous mycobacteria (NTM) are found in soil, water, and a variety of animals including invertebrates, reptiles, birds, and mammals (51). These species form biofilms, can live in environments with low nutrient content, and are easily aerosolized due to their hydrophobicity (51). Several species of NTM can survive phagocytosis by predatory amoebae and survive within them; this is potentially the origin of the ability of

M. tuberculosis to live intracellularly (51). NTM are not obligate parasites of humans, but many species are opportunistic pathogens (52). Immunodeficiency increases the risk of mycobacterial infection (52). NTM infections are typically acquired from the environment and not spread from person to person (44). In water treatment plants, NTM are 10 to 100 times more resistant than *E. coli* to the disinfectants chlorine, chloramine, chlorine dioxide, and ozone (52). NTM pulmonary infections have been increasing by about 5 to 10% annually for the last 20 years (52).

Some NTM, especially *M. marinum*, *M. fortuitum*, and *M. chelonae*, are aquatic zoonoses, causing systemic granulomatous inflammation in fishes that is currently untreatable (1). Mycobacteriosis in fish is problematic in aquaculture, industrial fishing, and ornamental aquaria. Humans can contract infections via contaminated water or fish. Human infection usually remains in the extremities, although it can cause deeper infections or become disseminated, especially in immunocompromised individuals. Phages of *M. fortuitum* (53, 54) and *M. chelonae* (55) have been isolated, although there are no reports of these being used in phage therapy or biocontrol. Prophages have been induced in *M. marinum* (2) and *M. fortuitum* (56).

CHAPTER 2

THE SEARCH FOR AQUATIC MYCOBACTERIOPHAGES

2.1 INTRODUCTION

Metagenomics and microscopy studies indicate that bacteriophages (phages) are significantly more diverse than previously believed (22). Culture techniques provide information about a phage's lifestyle, host range, plaque morphology, and other characteristics, and can produce a high titer lysate for full genome sequencing and electron microscopy, but can be biased towards phages amenable to the techniques used (22). Difficulties have been encountered in the isolation of phages capable of infecting *Clostridium difficile* (35) and *Bifidobacterium spp*. (36, 37). Some phages require a lower agar concentration (39) or the addition of glycine (40), antibiotics (41), or mucins (42) in order to form visible plaques. Larger phages will not be isolated if a typical filtration step is included (34). It is clear that some phages require specific conditions to complete a lytic infection cycle.

Mycobacteria are acid-fast bacilli with a characteristic mycolic acid-rich cell envelope (44). Almost 200 species are found in the genus *Mycobacterium*, including pathogens of humans and non-human animals (44). Phages infecting *M. tuberculosis*, *M. bovis, M. avium, and M. ulcerans* have been isolated and proposed as a treatment for infections for these pathogens (47). MuMC, *M. fortuitum*, and *M. chelonae* are aquatic zoonoses, causing deadly systemic illness in fishes which can spread to humans in contact with contaminated animals (1). Phages have been isolated that infect *M. fortuitum* (53, 54) and *M. chelonae* (55). Prophages have reportedly been induced

from *M. marinum* (2) and *M. fortuitum* (56). However, there are still very few known phages for these species.

The typical methods of phage isolation are direct plating, enrichment, and prophage induction (57). Direct plating involves using the environmental sample in a plaque assay without enrichment, while enrichment involves incubating the sample with the host of interest prior to phage detection (57). Environmental samples are selected based on where one would expect to find the host of interest; phages can usually be found in any environmental source, but not every sample will contain a phage infecting the host of interest (57). Concentration techniques may need to be used for samples with a low viral titer; polyethylene glycol (PEG) precipitation is one of the most common methods of phage concentration due to its simplicity and effectiveness (58). Detection of phages after enrichment is usually accomplished by a plaque assay (application of a filtrate of the sample or enrichment on a lawn of the bacterium of interest), with either undiluted sample or a dilution series (57). A specialized technique that is used for evolutionary and ecological research but can also be used to develop phages capable of infecting a particular host, is host range expansion (59). This technique involves sequential or concurrent culture of a phage with more than one bacterial species or strain in order to broaden the host range of the phage (59). Another method of phage isolation, often used to gain insight into the history of temperate phage infection of a bacterium, is prophage induction (18). Prophages are phages that have inserted their genetic sequence into the host genome; stress conditions, such as ultraviolet (UV) light or the DNA-damaging chemical mitomycin C (MMC), can cause the prophages to

resume a lytic cycle (induction) and the phages can then be isolated from the medium (18).

The aim of the present study was to isolate phages infecti*ng MuMC, M. fortuitum*, and *M. chelonae*, using both enrichment of environmental samples and induction of prophages from bacterial genomes. Many techniques were attempted, so that optimal conditions for phage isolation in these species could be established. As phages become more widely used in biotechnology, medicine, and other fields, more focus must be placed on determining the most effective methods of phage discovery for less-commonly isolated phage types.

2.2 MATERIALS AND METHODS

A brief summary of the experiments performed in the present study is listed in Table 2.

TABLE 2 Summary of experiments

Experiment	Method	Detection
Enrichment 1	Enrichment in flasks	Plaque assay
Enrichment 2	Enrichment in tubes	Plaque assay
Enrichment 3	Directed evolution (passage sample though multiple enrichments with increasing proportion of target bacterium)	Plaque assay
Enrichment 4	Enrichment in 6-well plates	Plaque assay and turbidity assay
Enrichment 5	PEG precipitation (concentration of viral fraction of sample)	Plaque assay and turbidity assay
Induction 1	Prophage induction with UV, heat, and MMC in tubes	Plaque assay
Induction 2	Prophage induction with MMC in tubes	Plaque assay
Induction 3	Prophage induction with MMC in 96-well plates	Plaque assay

Bacterial Cultures

Media was prepared by combining 13.35 mL Middlebrook 7H9 broth (7H9), 10 mL albumin dexrose supplement (AD) (for *M. smegmatis*) or oleic albumin dextrose catalase supplement (OADC) (for *M. chelonae*, *M. fortuitum*, and *M. marinum*), 1 mL 100 mM $CaCl_2$, 100 µL 50 mg/mL carbenicillin (CB), and 100 µL 10 mg/mL cyclohexamide (CHX). To prevent clumping, cultures were grown with 0.05% Tween 80 (polyoxyethylene sorbitan monooleate). The cultures were incubated at 37°C, shaking at 140 rpm, until reaching approximately 10^6 CFU/mL (estimated by turbidity, then confirmed by colony count). Because Tween can inhibit phage attachment, cultures were centrifuged at 2,000 x *g* for 10 minutes to pellet the bacteria, the supernatant was removed, and the pellet was resuspended in fresh medium (without Tween) before use.

Procedure Sources

Flask-based enrichments were based on the enrichment procedure in the Phage Discovery Guide (25). Multi-well plate methods are based on the methods described in Chapter 3: Improving High-Throughput Techniques for Bacteriophage Discovery in Multi-Well Plates. The host range expansion experiment is a modification of the procedure described by Burrowes and colleagues (60). The PEG precipitation procedure is based on the procedure from Antibody Design Labs (61). The prophage induction experiments were based on the work of Pettersson and colleagues (MMC and heat stress) (62), Ho and colleagues (MMC and heat stress) (63), and Fusco and colleagues (UV light) (64). These procedures are briefly described below, with modifications noted.

Statistical Analysis

For 96-well plate turbidity assays with dilution series, SPSS (Statistical Package for the Social Sciences) was used to perform a one-way Analysis of Variance (ANOVA) on each time point; a one-tailed Dunnett's test was used for post-hoc testing (65). For 96-well plate turbidity screenings (undiluted phage) a t-test was used. A p-value of less than 0.05 was considered statistically significant. For 96-well plate turbidity assays using undiluted samples, a statistically significant difference ($p<0.05$) in absorbance between the control group and the experimental group on at least three consecutive days was considered a positive result. For 96-well plate turbidity assays using dilution series, a statistically significant difference ($p<0.05$) in absorbance between the control group and at least three dilution groups on at least three consecutive days was considered a positive result.

Enrichment Experiment 1

Hosts: *M. chelonae* M3, *M. fortuitum* M5, and *M. marinum* ATCC927.

Phage sample: solid material and water from the filter of a warm freshwater ornamental aquarium in an office at Old Dominion University in March 2016.

The flask-based enrichment was a modification of the method in the Phage Discovery Guide (25). Bacterial cultures of each of the strains were added to a 250 mL flask at a volume of 5 mL with 40 mL 7H9 broth, 5 mL OADC, 0.5 mL 100 mM $CaCl_2$, and 1 mL of the aquarium sample. The flasks were incubated at 30°C, shaking at 140 rpm, for one week. Then, 4 mL of each enrichment was transferred to a clean 15 mL

tube, which was centrifuged at 2,000 x *g* for 10 minutes. The supernatant was sterilized with 0.2 µm filters and undiluted filtrate was plated by the agar overlay method, using both 0.2% and 0.4% agar, on the respective bacterial strains. The plates were incubated at 30°C.

Enrichment Experiment 2

Hosts: *M. chelonae* M3, *M. chelonae* M324-818, *M. fortuitum* M5, *M. fortuitum* M6, *M. marinum* ATCC927, and *M. marinum* M30-01.

Phage samples: four water samples from the Virginia Institute of Marine Sciences (VIMS) (Gloucester Point, Virginia, USA) in June 2016 - water and filter from uninfected aquarium, water and filter from infected aquarium, water and biological filtration ball from uninfected aquarium, water and biological filtration ball from infected aquarium.

The flask-based enrichment was a modification of the method in the Phage Discovery Guide (25). Bacterial cultures of each of the strains were added to four 5 mL tubes at a volume of 400 µL. To each of the tubes was added 3.2 mL 7H9 broth, 400 µL OADC, 40 µL 100 mM $CaCl_2$, and 50 µL of an aquarium sample (the four samples were added to each strain's four tubes, respectively). The tubes were incubated at 30°C, shaking at 120 rpm, for one week. Then, the enrichments were centrifuged at 2,000 x *g* for 10 minutes. Each supernatant was sterilized with 0.2 µm filters and plated by the agar overlay method, using both 0.2% and 0.4% agar, on the respective bacterial strains. The plates were incubated at 30°C for seven days.

Enrichment Experiment 3 (Directed Evolution/Host Range Expansion)

Hosts: *M. smegmatis* mc^{2}155, *M. fortuitum* M5, and *M. marinum* M.

Phage samples: soil from ODU campus collected in November 2016; pooled four soil samples from 1. soil under pine tree near pond (shade), 2. garden soil next to Orchid Observatory (sunny), 3. soil under Norway Maple (partly shaded), and 4. soil under azalea bushes near Orchid Observatory (sunny).

This experiment is a modification of the procedure described by Burrowes and colleagues (60), performed in flasks similarly to the enrichment procedure described by Poxleitner and colleagues (25). The expansion of host range beyond the strain level is usually a multi-step process with a high failure rate (59), so *M. smegmatis* concentrations in the enrichment solution were gradually decreased at each passage in order to allow the phages a means of multiplication (to avoid extinction of the phage), while still applying evolutionary pressure in the form of increasingly limited hosts.

In a 250 mL flask, 1 g soil, 21.5 mL 7H9 broth, 3 mL AD supplement, 0.5 mL 100 mM $CaCl_2$, and 5 mL *M. smegmatis* mc^{2}155 were combined and incubated at 37°C, shaking at 140 rpm for one day. The enrichment was centrifuged at 2,000 x *g* for 10 minutes and then sterilized with 0.2 µm filters and spotted on 0.2% and 0.4% agar overlays with *M. smegmatis* mc^{2}155 and with *M. marinum* M. The next enrichments were set up as follows in 250 mL flasks: 21.5 mL 7H9 broth, 3 mL OADC supplement, 0.5 mL 100 mM $CaCl_2$, 50 µL of the filtrate from the first enrichment, 1 mL *M. smegmatis* mc^{2}155 culture, and 4 mL of another bacterial culture (*M. marinum* M for enrichment A and *M. fortuitum* M5 for enrichment B). The two enrichments were incubated at 30°C, shaking at 140 rpm for three days. The enrichments were

centrifuged at 2,000 x *g* for 10 minutes, sterilized with 0.2 µm filters, serially diluted, and spotted on 0.2% and 0.4% agar overlays with *M. smegmatis* mc^2155 (both enrichments), *M. marinum* M (enrichment A), and *M. fortuitum* M5 (enrichment B). The third enrichments were performed in the same manner as the second, but with 0.5 mL *M. smegmatis* mc^2155 and 4.5 mL of *M. marinum* M (enrichment A)/*M. fortuitum* M5 (enrichment B). The fourth enrichments used 250 µL *M. smegmatis* mc^2155 and 4.75 mL of *M. marinum* M (enrichment A)/*M. fortuitum* M5 (enrichment B). The fifth enrichments used 100 µL *M. smegmatis* mc^2155 and 4.9 mL of *M. marinum* M (enrichment A)/*M. fortuitum* M5 (enrichment B). The sixth enrichments used 50 µL *M. smegmatis* mc^2155 and 4.95 mL of *M. marinum* M (enrichment A)/*M. fortuitum* M5 (enrichment B).

Enrichment Experiment 4

Hosts: *M. fortuitum* M5, *M. marinum* ATCC927, and *M. marinum* M

Phage samples: nine samples from public display aquaria the Virginia Aquarium (Virginia Beach, Virginia, USA) in August 2018: 1. shark tank, skimmer box; 2. turtles tank, skimmer box; 3. Red Sea tank, skimmer box and perimeter; 4. seahorses tank, biotower; 5. bonefish tank, biotower; 6. seagrasses tank, skimmer box; 7. crocodiles tank, skimmer box; 8. Asian turtles tank, skimmer box; 9. rays touch tank, biotower.

Nine water samples were obtained from filters, skimmer boxes, and biotowers in active aquaria at the Virginia Aquarium; the surfaces of each filter or skimmer box were swabbed and the swab immersed in the water sample from the same source in order to collect both planktonic viruses and those that may be contained within bacterial biofilms.

The 6-well plate enrichment procedure (described in in Chapter 3: Improving High-Throughput Techniques for Bacteriophage Discovery in Multi-Well Plates) was performed for each sample on each of the three strains and incubated at 30°C, shaking at 120 rpm, for five days. The enrichments were centrifuged at 2,000 x *g* for 10 minutes, sterilized with 0.2 µm filters, and spotted on 0.2% and 0.4% agar overlays with their respective hosts. Turbidity assays in 96-well plates were performed with undiluted samples (in triplicate, with controls).

Enrichment Experiment 5 (PEG Precipitation)

Hosts: *M. smegmatis* mc^2155, *M. chelonae* M3, *M. chelonae* M324-818, *M. fortuitum* M5, *M. fortuitum* M6, and *M. marinum* strains ATCC927, ATCC11564, KST214, M, and M2.

Phage samples: twelve water samples from active aquaria at a pet store in Virginia Beach, VA in February 2020. Samples 1-4: from warm saltwater tanks (1. horseshoe crab and cleaner shrimp tank, 2. lionfish tank, 3. clownfish tank, 4. corals tank), samples 5-8: from warm brackish water tanks (5. large scat tank, 6. fugu tank, 7. spotted puffer tank, 8. knight goby tank), samples 9-11: from warm freshwater tanks (9. gold severum tank, 10. cichlid tank, 11. goldfish tank), sample 12: from a cold freshwater koi tank.

Three pooled samples were created. The salt water pooled sample contained 10 mL from each of samples 1-4. The brackish water pooled sample contained 10 mL from each of samples 5-8. The fresh water pooled sample contained 10 mL from each of samples 9-12. A PEG precipitation method was performed on each pooled sample (61). Briefly, each tube was centrifuged for 5 minutes at 13,000 x *g* and 35 mL of supernatant

was transferred to a new tube. PEG/NaCl solution (9 mL) was added to each tube and inverted to mix. The tubes were chilled on ice for 1.5 hours, then centrifuged at 13,000 x *g*. The supernatant was discarded, the tubes were centrifuged for 5 more minutes, and the supernatant discarded again. The pellets were each resuspended in 4 mL TBS and vortexed. The tubes were chilled on ice for one hour, vortexed, and centrifuged at 13,000 x *g* for 5 minutes. The supernatant was collected and filter sterilized with 0.2 µm filters.

A spot test was performed with a dilution series of each filtered sample on each strain with 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*). A 96-well plate detection assay (described in in Chapter 3: Improving High-Throughput Techniques for Bacteriophage Discovery in Multi-Well Plates) was also performed with each filtered sample (and control wells for each strain) in triplicate with dilution series of the filtered samples.

Prophage Induction Experiment 1

Hosts: *M. marinum* strains ATCC927, ATCC11564, C7, KST214, M2, M4, M11, M12, M13, M30-01, M59, M324-958, R171, and Rp72a. Resulting filtrates were also plated on *M. smegmatis* mc^2155, *M. chelonae* M3, *M. fortuitum* M5, and *M. fortuitum* M6.

Conditions: auto-induction, heat, UV, mitomycin C (MMC)

These experiments were based on the work of Pettersson and colleagues (62), Ho and colleagues (63), and Fusco and colleagues (64).

Auto-induction/control: A 1 mL sample from each bacterial culture was centrifuged at 2,000 x *g* for 10 minutes, then sterilized with 0.2 µm filters. A spot test

was performed on each strain with 0.2% and 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*).

Heat: For each strain, 1 mL of bacterial culture was removed to a new 5 mL tube with 1 mL fresh medium. The tubes were incubated at 36°C, shaking at 140 rpm, for three days. From each tube, 1 mL was centrifuged at 2,000 x *g* for 10 minutes, then sterilized with 0.2 μm filters. A spot test was performed on each strain with 0.2% and 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*).

UV: The UV light used for some of the prophage induction experiments was a 254 nm light from a Philips 30 W T8 bulb in a biological safety cabinet at a distance of 1 meter (intensity 100 μW/cm^2). For each strain, 1 mL of bacterial culture was removed to a new 5 mL tube with 1 mL fresh medium. The tubes were incubated at 30°C, shaking at 140 rpm, for three days. On the starting day and the two subsequent days, the tubes were placed under a UV light for 15 minutes. After three days, 1 mL from each tube was centrifuged at 2,000 x *g* for 10 minutes, then sterilized with 0.2 μm filters. A spot test was performed on each strain with 0.2% and 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*).

MMC: MMC (Santa Cruz Biotechnology, Dallas, TX, USA) was prepared by adding 2 mL double-distilled water to the 2 mg MMC powder in the vial. The solution was mixed until the powder dissolved and filter sterilized with a 0.2 μm filter to produce a 1 mg/mL MMC stock solution. For each strain, 1 mL of bacterial culture was added to each of four 5 mL tubes and 1 mL fresh medium added to each. The concentrations of MMC that were added to the four tubes were: 0.25 μg/mL, 0.5 μg/mL, 1.0 μg/mL, and 1.5 μg/mL. The tubes were incubated at 30°C, shaking at 140 rpm, for three days. For

each strain, 1 mL from each MMC level was pooled together and centrifuged at 2,000 x *g* for 10 minutes, then sterilized with 0.2 µm filters. A spot test was performed on each strain with 0.2% and 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*). If any visible plaques were observed, spot tests for the individual dilutions would be performed to deduce which dilutions resulted in successful prophage induction.

Prophage Induction Experiment 2

Host: *M. marinum* M. Resulting filtrates were plated on: *M. smegmatis* mc^2155, *M. chelonae* M3, *M. fortuitum* M5, and *M. marinum* strains Davis, M, M12, and M30-01.

Condition: MMC

This experiment was based on the work of Pettersson and colleagues (62) and Ho and colleagues (63).

MMC (Santa Cruz Biotechnology, Dallas, TX, USA) was prepared by adding 2 mL double-distilled water to the 2 mg MMC powder in the vial. The solution was mixed until the powder dissolved and filter sterilized with a 0.2 µm filter to produce a 1 mg/mL MMC stock solution.

In each of six 50 mL tubes, 2.5 mL of *M. marinum* M was combined with 2.5 mL fresh medium. To each tube, a different concentration of MMC was added: 0 µg/mL (control), 1 µg/mL, 3 µg/mL, and 5 µg/mL, 10 µg/mL, and 20 µg/mL. The tubes were incubated at 30°C, shaking at 140 rpm, for nine days. For each tube on days 3, 5, 7, and 9, 1 mL was removed and centrifuged at 2,000 x *g* for 10 minutes, then sterilized

with 0.2 µm filters. A spot test was performed on each strain with 0.2% and 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*).

Prophage Induction Experiment 3

Hosts: *M. chelonae* M3 and *M. marinum* M. Resulting filtrates were plated on *M. smegmatis* mc^2155, *M. chelonae* M3, *M. fortuitum* M5, and *M. marinum* strains ATCC927, M, M13.

Condition: MMC

The prophage induction methods were based on the work of Pettersson and colleagues (62) and Ho and colleagues (63).

MMC (Santa Cruz Biotechnology, Dallas, TX, USA) was prepared by adding 2 mL double-distilled water to the 2 mg MMC powder in the vial. The solution was mixed until the powder dissolved and filter sterilized with a 0.2 µm filter to produce a 1 mg/mL MMC stock solution.

In a 96-well plate, 45 wells were used per strain. For each strain, five different concentrations of MMC were used: 0 µg/mL (control), 1 µg/mL, 3 µg/mL, 5 µg/mL, and 10 µg/mL (nine wells were used per MMC concentration), which were added to the typical solution used for a 96-well plate enrichment procedure (see section 3.2). The 96-well plate was incubated at 30°C, shaking at 90 rpm, for three days. After one day, three wells per condition were pooled, centrifuged at 2,000 x *g* for 10 minutes, and sterilized with 0.2 µm filters. This was also done after two days and after three days. A spot test was performed with each filtered sample on each strain with 0.4% agar overlays, which were incubated at 30°C (37°C for *M. smegmatis*).

2.3 RESULTS

No plaques were seen for enrichment experiments 1 and 2. In enrichment experiment 3, the plaque assay of the second round of enrichment A showed areas of clearing on *M. marinum* M for both 0.2% and 0.4% agar overlays (one large area of clearing on the 10^{0}-10^{-2} dilutions and possible individual plaques on the 10^{-3} dilutions). These results were not reproduced when touching a wooden stick to the plaques and transferring to a new plate or by full plate flooding; re-plating of the original sample was also unsuccessful. In enrichment experiment 4, samples 2, 3, and 9 showed possible areas of light clearing on *M. marinum* ATCC927 and sample 6 showed possible light clearing on *M. marinum* M (all areas of clearing were only on most concentrated phage sample and no individual plaques were seen).

In the 96-well plate turbidity assays for enrichment experiment 4, there were positive results (see Statistical Analysis in section 2.2 for a definition of positive results) for samples 2 and 3 paired with *M. marinum* ATCC927 (Fig. 2). In enrichment experiment 5, no plaques were seen; however, the 96-well turbidity assay had positive results for *M. marinum* KST214 with the freshwater, brackish water, and saltwater samples, *M. chelonae* M324-818 had positive results for the brackish water sample, and *M. fortuitum* M5 had positive results for the saltwater sample (Fig. 3).

No plaques were seen for prophage induction experiments 1, 2, and 3. MMC was associated with a reduction in bacterial growth in two representative strains, *M. marinum* M and *M. chelonae* M3 (Table 3).

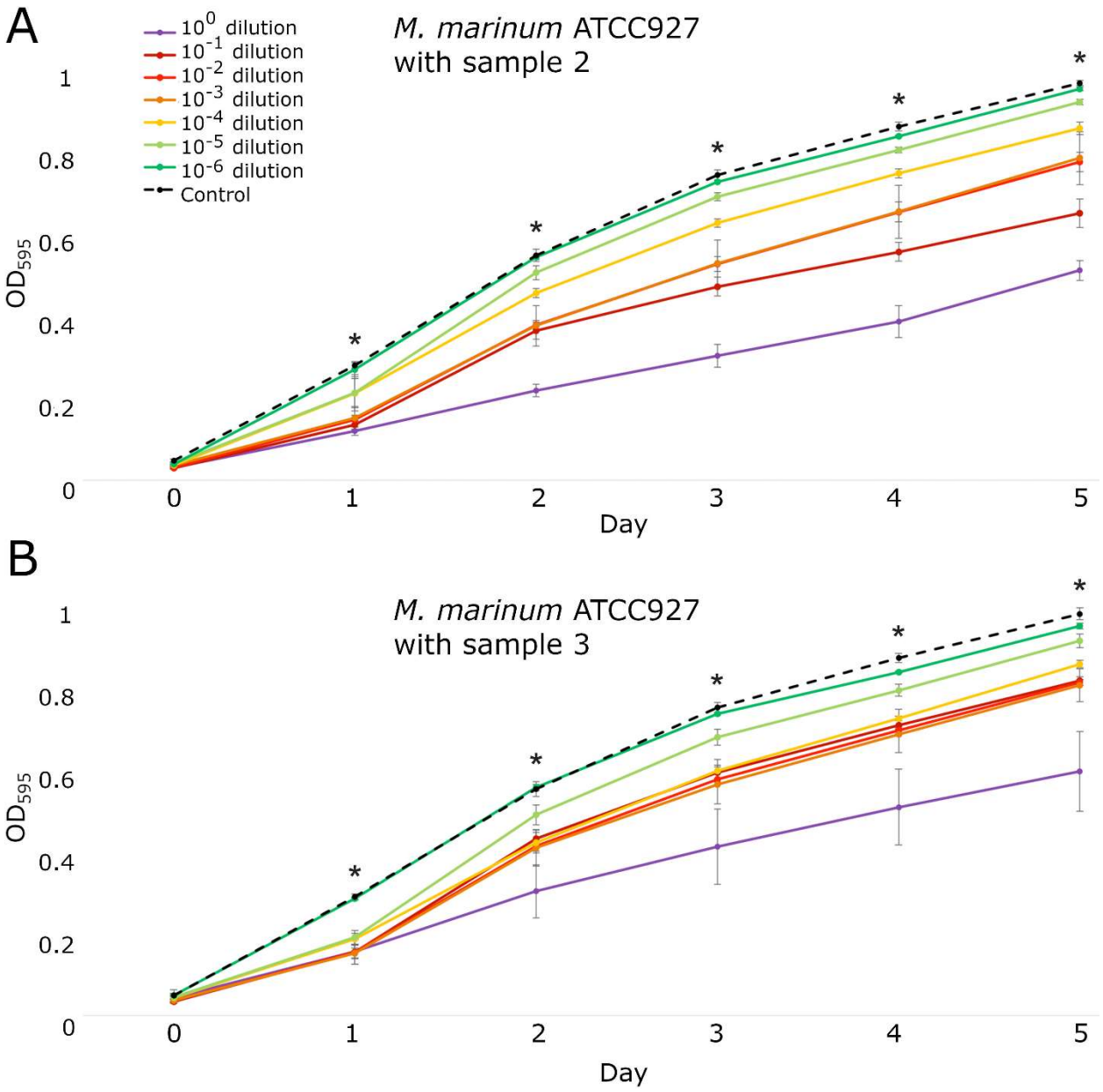

FIG 2 Enrichment 4 positive turbidity assay results. (A) *M. marinum* ATCC927 with sample 2, (B) *M. marinum* ATCC927 with sample 3. Error bars represent +/- 1 standard deviation of triplicate wells. Asterisks indicate days in which at least three dilution levels had a statistically significant difference from the control.

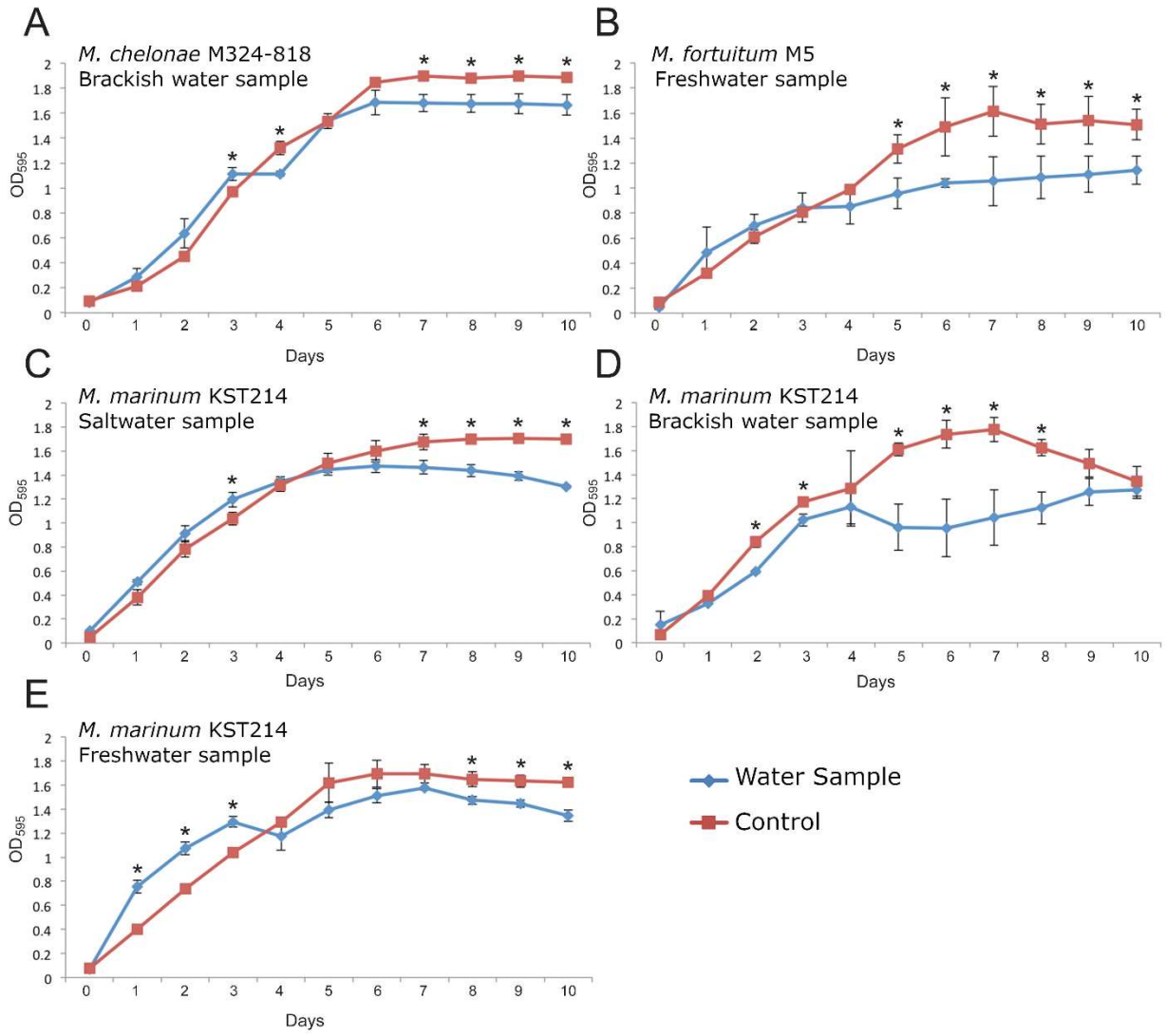

FIG 3 Enrichment 5 positive turbidity assay results. (A) *M. chelonae* M324-818 with brackish water sample, (B) *M. fortuitum* M5 with freshwater sample, (C) *M. marinum* KST214 with saltwater sample, (D) *M. marinum* KST214 with brackish water sample, E: *M. marinum* KST214 with freshwater sample. Error bars represent +/- 1 standard deviation of triplicate wells. Asterisks indicate days in which the experimental group had a statistically significant difference from the control.

TABLE 3 MMC concentration vs OD_{595} during three days of incubation for two strains representing greater (*M. chelonae* M3) and lesser (*M. marinum* M) MMC sensitivity. A gray background represents that the difference between the experimental and control group has a p value < 0.05 (one way ANOVA with post-hoc Dunnett test)

Day	MMC Concentration (µg/mL)	Average OD_{595} *M. chelonae* M3	Average OD_{595} *M. marinum* M
0	10	0.57	0.75
	5	0.59	0.76
	3	0.60	0.59
	1	0.62	0.65
	0 (Control)	0.60	0.55
1	10	0.57	1.00
	5	0.60	1.02
	3	0.61	0.98
	1	0.65	1.02
	0 (Control)	0.67	1.05
2	10	0.59	1.09
	5	0.62	1.12
	3	0.63	1.11
	1	0.66	1.11
	0 (Control)	0.69	1.16
3	10	0.65	1.14
	5	0.65	1.18
	3	0.66	1.18
	1	0.68	1.16
	0 (Control)	0.73	1.22

2.4 DISCUSSION

No phages were definitively found to produce individual plaques on the tested bacterial strains. In enrichment experiment 3, individual plaques were seen but could not be replicated by three attempted methods. The initial results may have been the result of contamination of a different bacterial strain that was susceptible to the phage, or the phage may have been unstable and quickly lost its ability to complete a productive infection cycle. In enrichment experiment 4, light regions of clearing were

only seen in the most concentrated samples and no individual plaques were seen. This phenomenon is usually attributed to lysis from without, in which a high density of phages adsorb to the bacterial cell envelope, resulting in decreased bacterial growth in the locations where a high titer of phage was applied but no formation of individual plaques (66). These observations can also be the result of abortive infection (in which bacterial death inactivates infecting phages) or harmful compounds in the phage solution (66).

Many enrichment techniques with a variety of environmental sources were attempted. Enrichment experiment 4, using samples from the Virginia Aquarium, produced some positive results which may be indicative of some level of phage-host interaction (Fig. 2). These were some of the same sample-host combinations that produced faint areas of clearing but no formation of individual plaques in the plaque assay; therefore, the turbidity assay results may be showing the same partial phage infection that is indicated by the plaque assay results. It is also possible that the phage-host interactions occurring in the liquid medium are different from the interactions occurring on the semisolid medium. Unfortunately, it is still unknown how to distinguish a productive phage infection from a partial phage infection in turbidity assays. Similarly, enrichment experiment 5 resulted in some positive turbidity assays (see Statistical Analysis in section 2.2 for a definition of positive results), but the plaque assays displayed no indication of decreased bacterial growth (Fig. 3). Monitoring turbidity assays for longer time periods sometimes led to positive outcomes on days six or later (Fig. 3), indicating that in some cases bacterial growth inhibition may be visible only when overcrowding and nutrient deprivation increase stress on the bacterial cells.

No results were observed for the prophage induction experiments. In 1969, Bönicke reported auto-induction of prophages from a strain of *M. marinum* and was able to infect other *M. marinum* strains with them (2). Although in Prophage Induction Experiment 1 a similar experiment with auto-induction was performed (in addition to other induction methods in Prophage Induction Experiments 1, 2, and 3), it is possible that the strains used by Bönicke had a greater tendency towards auto-induction or that the indicator strains were more susceptible to the phage. Although large prophages are found in some strains of *M. marinum*, including many used in these studies (Chapter 4: Phylogeny and Evolution of Prophages in 49 Genomes of *Mycobacterium marinum*), the only way to confirm the intactness of a prophage is to demonstrate successful induction (67). It is possible that the prophages within the bacterial strains used in the present studies are not fully intact, or that susceptible hosts have not yet been identified.

The results of the present study highlight the difficulty of isolating phages for some bacterial host species. The plaque assays were uninformative, unreplicable, or did not result in individual plaques. Turbidity assays sometimes indicated possible phage-host interactions, but the results were often subtle, and it is not possible at this time to interpret fully the significance of these results. Interestingly, positive turbidity assay results were sometimes obtained for host-sample pairings that had not produced any observable effects on the plaque assay, indicating that certain phage-host interactions may be detectible in a liquid medium but undetectable on a semisolid medium. In the future, it is important to work towards a greater understanding of phage-host interactions in a variety of media types and how best to measure and interpret the results.

CHAPTER 3

IMPROVING HIGH-THROUGHPUT TECHNIQUES FOR BACTERIOPHAGE DISCOVERY IN MULTI-WELL PLATES

3.1 INTRODUCTION

Interest in bacteriophages (phages) has been growing as their numerous applications become more fully appreciated (68). Phages can be used for medical applications (phage therapy), to kill bacteria in natural and man-made environments (biocontrol), for molecular biology techniques, and in other applications (11).

Phage therapy has been growing in popularity due both to the rise in antibiotic resistant bacteria, and because phages have several advantages compared to chemical antibiotics. Beneficial bacteria are not harmed if phages with narrow host ranges are selected, no serious adverse effects have been reported in clinical trials, using a cocktail of phages can reduce problems with resistance, and phages are often easier to isolate than new chemical antibiotics (31). Phage cocktails can be tailored to the exact strain of bacteria responsible for a patient's infection by screening the strain against a phage library; the cocktail can be adjusted if the pathogen evolves (69). Phage therapy has a long history in some countries, but is considered an investigational drug in others, such as the United States. However, there are many proponents of phage therapy who are working towards making this treatment available to greater numbers of patients (32). Inexpensive, rapid, high-throughput methods will be necessary to meet any increasing demand for therapeutic phages.

Standard phage discovery techniques involve two major steps. The first, enrichment, consists of incubating an environmental sample with the host bacterium, then filtering the solution to remove bacteria and transferring the resulting filtrate to a lawn of the host bacterium (plaque assay). If successful, plaques (areas of clearing) will appear, which indicate lysis of the bacteria by the phages (25). The second step, purification, usually involves selecting one plaque from the plaque assay to use as the sample for another plaque assay and repeating this for several passages until plaque morphology suggests that a pure culture of phage has been attained. Consistent plaque morphology is often used as a measure of the purity of the sample; nucleic acid sequencing, RAPD-PCR, and RFLP analysis could help to determine purity but are rarely used for this purpose (25, 70). Enrichment is usually performed in a liquid medium with a relatively large volume (~ 50 mL) of broth in a flask; purification is usually performed on a semisolid medium (double layer agar overlay on Petri plates) (25). Phages destined for phage therapy must then undergo thorough screening procedures for characteristics including specificity, toxin production, and whether they are lytic or temperate (57). When traditional isolation techniques are performed on a large scale, the amount of materials, time, and space used by these procedures can inhibit rapid discovery.

Some alternatives to plaque assays have been published, but they are limited in their applications and are rarely used in phage discovery. Early in the history of phage research, the presence of phages was determined by an observation of decreased bacterial culture turbidity (71, 72), but this method is now uncommon except within research concerning viruses of eukaryotes (73) or the production of phage stock

solutions (74). Metagenomics studies are becoming more prevalent, but these give incomplete genetic data (75, 76). Polymerase chain reaction (PCR) can detect phages if portions of the phage genome are known (77). Antibody-based methods such as enzyme-linked immunosorbent assay (ELISA) and fluorescent dyes can be used to detect a phage, but an antibody to the phage must be available (78). It is also important for certain analyses (such as electron microscopy, full genome sequencing, and host range studies) that the phage is propagated to make a high titer stock solution, and many techniques do not accomplish this.

Multi-well plate methods can improve efficiency, save space, and reduce cost. Multi-well plates have already been used for a variety of techniques in phage research, including enrichment (79), host infection/ phage amplification, biofilm eradication experiments (30), prophage induction (80), directed evolution (81), and phage display (82). Agar can be poured into the wells of 96-well plates for high-throughput plaque assays, although this is less common than using liquid medium in multi-well plates (83). Bioluminescent reporter genes have been inserted into phages so that the phages can detect their host via luminescence readings in multi-well plates (84, 85). The first description of a turbidity-based phage detection method using serial dilutions of a sample in a liquid medium was given by Appelmans in 1921 (71). Subsequently, similar methods have been performed using either absorbance readings or live-dead colorimetric assays in multi-well plates (86-90). Despite advances in research methodology, much of phage research is still performed using traditional practices. Increasing the number of phage techniques that can be performed in multi-well plates

and optimizing current multi-well techniques could provide great benefit to phage discovery research.

Using liquid media instead of agar plates may also have the benefit of improving phage discovery for certain phages. Phages are diverse, not only in their morphology, but also in the properties required for host adsorption, growth, and propagation. Some phages will only form plaques, or will form larger plaques, with the addition of glycine (40) or sublethal levels of antibiotics (41). Phages of *Bifidobacterium* have only been discovered via PCR, electron microscopy, and the presence of prophages in their genomes; plaque assays do not result in observable plaques (36). The spatial structuring that a semisolid medium imposes leads to different selective pressures on a phage in agar than in broth; agar can be a barrier to phage diffusion, which can hinder adsorption to the host (39). In fact, some phages have been observed to only form plaques on media with a low agar concentration (34). Since a liquid medium is typically used in multi-well plates, a benefit of multi-well plates may be the isolation of a greater diversity of phages than can be discovered using agar-based techniques. A greater diversity of phages could be discovered if conditions could be adapted to the needs of specific phages.

Developing new phage isolation techniques may allow discovery of conditions that are best suited to different phages and in turn may improve isolation methods for aquatic phages, which are isolated at much lower rates compared to terrestrial phages (23). In fact, most of the aquatic phages that have been identified to date are only known through metagenomics data (75). An analysis of these metagenomic data indicates that there is a much greater diversity of phages in the environment than what

can be seen in currently isolated phages (75, 76, 91). Phages that have been isolated and propagated in the laboratory are not necessarily representative of the phages present in the environment; those that are more easily propagated using current techniques may be overrepresented (22).

The present studies utilize the saprophytic *Mycobacterium smegmatis* as a model organism, since optimal conditions for phage procurement have already been determined and much is known about phages of this organism (25, 27, 92). *M. smegmatis* is not usually pathogenic, but it is in the same genus as many important pathogens of humans, livestock, and other animals, making it an ideal model organism (27, 93). The most well-known human pathogens in the genus *Mycobacterium* include *M. tuberculosis* and *M. leprae*, but it is also noteworthy that the incidence and prevalence of nontuberculous mycobacterial infection in humans has been increasing in recent decades (94). Phages and their components have been investigated as a treatment for many different mycobacterial infections (47, 48), as mycobacteria have a specialized cell wall that allows them to endure many stressors including several antibiotics (44). Recently, phage derivatives have been developed using genome engineering and used in combination with phage cocktails against *M. abscessus* (48). Phage-based assays are also used in tuberculosis diagnostics and drug resistance testing (95).

In this work, pre-existing multi-well plate methods were adapted to mycobacteria and these methods were extended to purification of phages in liquid medium. Procedures for phage enrichment and detection in multi-well plates were optimized using a variety of environmental samples and previously purified phages, with *M.*

smegmatis mc^{2}155 as the host. A procedure was developed for the purification of phages in multi-well plates, which is not less effective or substantially more time consuming than the traditional methods. Hopefully, these methods will continue to be improved upon and streamlined, which may eventually be adapted to additional host species.

3.2 MATERIALS AND METHODS

Environmental Sample Collection

Soil samples were collected by filling 15 mL sterile centrifuge tubes with soil from three sites in southeastern Virginia, U.S.A.: Old Dominion University (ODU) campus in Norfolk, Virginia, the Great Dismal Swamp in Suffolk, Virginia, and a residential area in Norfolk, Virginia. Table 4 shows the locations of each soil sample collection.

Bacterial Strains and Purified Phages

The bacterial host used was *M. smegmatis* mc^{2}155. Cultures were grown by combining 13.35 mL Middlebrook 7H9 broth (7H9), 10 mL albumin dextrose (AD) supplement, 1 mL 100 mM $CaCl_2$, 100 µL 50 mg/mL carbenicillin (CB), and 100 µL 10 mg/mL cyclohexamide (CHX). To prevent clumping of bacteria, cultures were grown with 0.05% Tween 80 (polyoxyethylene sorbitan monooleate). The cultures were incubated at 37°C, shaking at 140 rpm, until reaching approximately 10^6 CFU/mL (estimated by turbidity, then confirmed by colony count), which took approximately three days. Because Tween can inhibit phage attachment, cultures were centrifuged at 2,000

x *g* for 10 minutes to pellet the bacteria, the supernatant was removed, and the pellet was resuspended in fresh medium (without Tween) before use.

The pre-purified phages used in this study had been isolated and purified by undergraduate students of the Phage Discovery course at ODU over the course of three years (2014-2016), using the methods in the Phage Discovery Guide (25). All phages were in the Order *Caudovirales*, Family *Siphoviridae*. Characteristics of these phages can be seen in Table 5. Lifestyle, cluster, and subcluster data were retrieved from the Actinobacteriophage Database at phagesdb.org, which based these designations on genetic data (92).

TABLE 4 Origins of soil samples used in this study

Sample	Coordinates
ODU-1	36.885286, -76.307138
ODU-2	36.885617, -76.306639
ODU-3	36.884934, -76.306941
GDS-1	36.741833, -76.379241
GDS-2	36.742435, -76.379780
GDS-3	36.728894, -76.386043
RES-1	36.869248, -76.219953
RES-2	36.869114, -76.219839
RES-3	36.869087, -76.219921

TABLE 5 Phages used in this study. Plaque size is based on growth at 37°C for 24 hours

Phage Name	Lifestyle	Plaque Size Range (mm)	Cluster	Subcluster	Location of Sample Collection
LordCommander	Temperate	0.25 - 1.00	A	A2	Norfolk, VA
BakterODU	Temperate	0.25 - 1.50	A	A4	Virginia Beach, VA
Exp626	Temperate	0.25 - 3.00	A	A4	Norfolk, VA
Jinger	Temperate	0.25 - 1.00	A	A4	Virginia Beach, VA
Rosepeake4	Temperate	0.25 - 1.50	A	A4	Chesapeake, VA
Goldilocks	Temperate	0.25 - 1.50	E	[none]	Norfolk, VA
CircleStick	Temperate	0.25 - 1.00	K	K1	Suffolk, VA
SwampThing	Temperate	0.25 - 1.00	K	K1	Suffolk, VA
SWerbenman	Temperate	0.25 - 1.50	K	K1	Suffolk, VA
Unicorn	Temperate	0.25 - 0.50	K	K6	Suffolk, VA
BobLoblaw	Lytic	0.50 - 1.50	M	M1	Norfolk, VA
Bricole	Lytic	0.50 - 1.50	M	M1	Virginia Beach, VA

Statistical Analysis

SPSS (Statistical Package for the Social Sciences) was used to perform a one-way Analysis of Variance (ANOVA) on each time point for the detection assays and purification assays. A one-tailed Dunnett's test was used for post-hoc testing (65). A p-value of less than 0.05 was considered statistically significant.

Enrichment in Multi-Well Plates

Enrichments were performed using multi-well plate methods in 6-well, 24-well, and 96-well plates. These were compared to enrichments from the same environmental samples performed using the traditional flask method (25).

Samples were collected according to the procedures described above (Environmental Sample Collection). Soil was mixed with the bacterial culture (approximately 10^6 CFU/mL) and the components described in Table 6, which is a

modification of the concentrations used by Poxleitner and colleagues in the Phage Discovery Guide (25). Flasks and plates were then incubated at 37°C, shaking at 140 rpm, for 48 hours. Following incubation, the fluid from each flask or well was centrifuged at 2,000 x *g* until the supernatant appeared transparent (20-60 minutes); the supernatant was then filtered through a 0.2 µm polyethersulfone filter (VWR International, Radnor, PA) to remove bacteria. Plaque assays were performed on the resulting filtrates in order to determine phage presence and titer. The presence of any plaques was considered a positive result and the titers of each positive result were compared.

TABLE 6 Volumes of components of each enrichment solution. Abbreviations: Middlebrook 7H9 broth (7H9), albumin dextrose supplement (AD), 100 mM $CaCl_2$ ($CaCl_2$), 50 mg/mL carbenicillin (CB), 10 mg/mL cyclohexamide (CHX)

Container Type	7H9	AD	$CaCl_2$	CB	CHX	Bacterial Culture	Soil
Flask	18.5 mL	2.5 mL	250 µL	15 µL	15 µL	1.25 mL	0.5 g
6-Well Plate	6 mL	720 µL	80 µL	8 µL	8 µL	400 µL	0.16 g
24-Well Plate	1.5 mL	180 µL	20 µL	2 µL	2 µL	100 µL	0.04 g
96-Well Plate	148 µL	20 µL	2 µL	0.2 µL	0.2 µL	10 µL	0.004 g

Phage Detection in Multi-Well Plates

Dilutions of pre-purified phage (20 µL per well of dilutions starting at 10^8 PFU/mL) mixed with host bacteria (10 µL per well of approximately 10^6 CFU/mL) were used to identify the lowest concentration of phage that could be detected by the 96-well plate

method. This is equivalent to approximately 1 x 10^4 bacterial cells per well mixed with a range of zero to 2 x 10^6 phages.

A plaque assay was performed on phage solutions one day prior to phage detection experiments so the appropriate concentrations could be estimated. Eight dilutions of each phage solution (beginning with 10^8 PFU/mL) were used along with a control (phage buffer), each performed in triplicate. Within each well of the 96-well plate, the components and volumes were the same as those used in the 96-well plates of the enrichment experiment, with the exception of 20 µL of the phage-containing solution being substituted for the soil sample. The absorbance at 595 nm was measured in a plate reader (BMG Labtech, Cary, NC) directly after mixing the solution and every 24 hours afterwards for a total of three days. Because it was found that some of the phages exhibit a rapid decrease in titer during storage, another plaque assay was performed at the time of starting the phage detection experiment and read the next day. Corrections were made to the starting concentrations in order to account for phage titers being less than expected based on the initial plaque assays. A positive result was defined as a statistically significant difference ($p<0.05$) in absorbance between the control group and any of the experimental groups.

Purification in Multi-Well Plates

Four pre-purified phages, each representing a different cluster (Table 5), were selected and mixed in equal parts (PFU) to form the starting phage mixture. These phages had been purified by undergraduate students according to the SEA-PHAGES protocols and were considered pure based on the traditional criteria of plaque

morphology and DNA sequencing (25). Serial dilutions of the phage mixture were combined with the host bacterium in wells of a 96-well plate in triplicate (similar to the procedure described above in Phage Detection in Multi-Well Plates). The starting bacterial concentration was 10^6 CFU/mL (estimated by turbidity, then confirmed by colony count) and the starting total phage concentration was 10^8 PFU/mL. Absorbance at 595 nm was measured immediately after combining the components and then every 24 hours for a total of three days. The most dilute group of wells that produced a noticeable decrease in turbidity from the control wells were pooled, filtered, diluted, and used for a new round of purification, performed in the same manner as the first. Biofilm formation was sometimes observed in the wells containing a mixture of phages, which was not usually a problem with the phage detection experiments involving a single phage type. This affected the accuracy of the OD_{595} readings and required correction in order to distinguish the phage-containing wells. Wells containing obvious biofilm formation were subject to visual examination to distinguish wells with an observable difference in turbidity from the control wells. Selection of wells sometimes had to be performed in a qualitative manner. In the future it would be beneficial to develop a means of reducing biofilm-related error that does not require resorting to qualitative methods. Ten rounds of purification were performed.

A traditional plaque assay was performed at each round of purification in order to determine the homogeneity of plaque morphology and the phage titer. Phage harvest was performed according to the method in the Phage Discovery Guide (25). Chromosomal DNA extraction was performed on the phage harvests using the PCI/SDS DNA extraction procedure on the Actinobacteriophage Database website (92).

The DNA was subjected to a RFLP with the restriction enzymes EcoRI and SexAI. For each reaction, 0.4 µL of SexAI (2 U), 0.2 µL of EcoRI-HF (4 U), 2 µL of 10X CutSmart Buffer, 1 µg of DNA, and molecular-grade water to 20 µL was combined and incubated overnight at 37°C. Heat inactivation was performed by incubating at 65°C for 20 minutes. A 1% agarose gel electrophoresis with GelStar Nucleic Acid Gel Stain (FMC BioProducts) was used to visualize the banding patterns on a D3-14 OWL gel electrophoresis system (Thermo Scientific), which were compared against 1 µL of 1 Kb+ Plus DNA Ladder (Thermo Fisher). The wells of the gel were each loaded with 5 µL of sample mixed with 1 µL of Gel Loading Dye Purple (New England BioLabs). The electrophoresis was run at 120 V for the first 60 minutes, followed by 80 minutes at 140 V, then visualized on a Universal Hood II Gel Doc System (Bio-Rad).

PCR reactions were also performed with a denaturation temperature of 95°C (30 seconds), an annealing temperature of 69°C for Bricole and Goldilocks primers, 65°C for CircleStick primers, and 70°C for Exp626 primers (45 seconds), and an extension temperature of 72°C (1 minute), for 35 cycles; each reaction contained 1 µM of each primer and 1X EconoTaq PLUS GREEN (Lucigen) mastermix (Table 7). The PCR reactions were run on a C1000 thermal cycler (Bio-Rad). To visualize the results of the PCR reactions, a 1.5% agarose gel with SYBR Safe gel stain (Thermo Fisher) was run on a horizontal electrophoresis system (Fisher Scientific) with a PowerPac HC (Bio-Rad) as the power source. Each PCR sample (5 µL per sample) was compared against 5 µL of Quick-Load Purple 100 bp DNA Ladder (New England BioLabs). The electrophoresis was run at 120 V for 40 minutes, then visualized on a Universal Hood II Gel Doc System (Bio-Rad).

TABLE 7 Primers used in the 96-well plate purification experiment

Target Phage	Primer Direction	Primer Sequence	Predicted Product Size	Target Gene
Bricole	Forward	5'-CTCGATGGTGTGGCAGAACT-3'	250 bp	Lysin B
	Reverse	5'-GTCATCTGGTGTGTCCGTGT-3'		
CircleStick	Forward	5'-ACGAATTGCGCCGGTTGCAC-3'	660 bp	Terminase
	Reverse	5'-ACCGCGGCCGAAACCTTTCT-3'		
Exp626	Forward	5'-CCCAAGGCAGTGTAATCGGT-3'	478 bp	Lysin B
	Reverse	5'-AGGTGCTTGACTGCGTACTC-3'		
Goldilocks	Forward	5'-GCTCTCCGTCTTGGTGTGAA-3'	321 bp	Lysin B
	Reverse	5'-CGACCCGAACAGGATTTCCA-3'		

3.3 RESULTS

Enrichment

Table 8 lists the concentration of phages detected in each enrichment solution at the end of the incubation period. Phages were detected in all three traditional flask enrichments. The three soil samples from the Great Dismal Swamp (GDS) produced detectable phage in all enrichments, even those performed in the 96-well plates, which are only able to accommodate a very small quantity of soil. All enrichment types for GDS samples produced a similar ending concentration of phages. The soil samples from the university campus (ODU) and residential area (RES) were, in most cases, able to produce detectable phage in 6-well plates, but not in 24- or 96-well plates. Direct plating of each soil sample was performed according to the procedure in the Phage Discovery Guide, but no plaques were observed (25).

Detection

The twelve previously isolated and purified phages used in the multi-well plate detection assay all produced a statistically significant difference in optical density (OD_{595}) when compared to the control wells in at least two of three replicates (Fig. 4). The phages varied in two ways: 1. in the degree to which the OD_{595} of the experimental groups differed from that of the control group and 2. in the length of time required before a difference between the OD_{595} of the experimental groups and that of the control group could be observed.

TABLE 8 Ending concentration of phages (in plaque-forming units per mL) after enrichment for each sample.

Name	Flask	6-Well Plate	24-Well Plate	96-Well Plate
ODU-1	4.2×10^6	2.0×10^4	0	0
ODU-2	6.4×10^7	4.2×10^5	0	0
ODU-3	9.8×10^7	1.5×10^8	1.6×10^3	0
GDS-1	1.2×10^7	5.5×10^7	9.5×10^8	5.5×10^7
GDS-2	4.5×10^7	7.5×10^8	1.1×10^8	1.2×10^7
GDS-3	7.0×10^8	8.0×10^8	1.1×10^9	1.3×10^7
RES-1	1.2×10^3	0	0	0
RES-2	1.2×10^8	9.6×10^8	0	0
RES-3	7.4×10^{10}	6.8×10^9	1.4×10^3	0

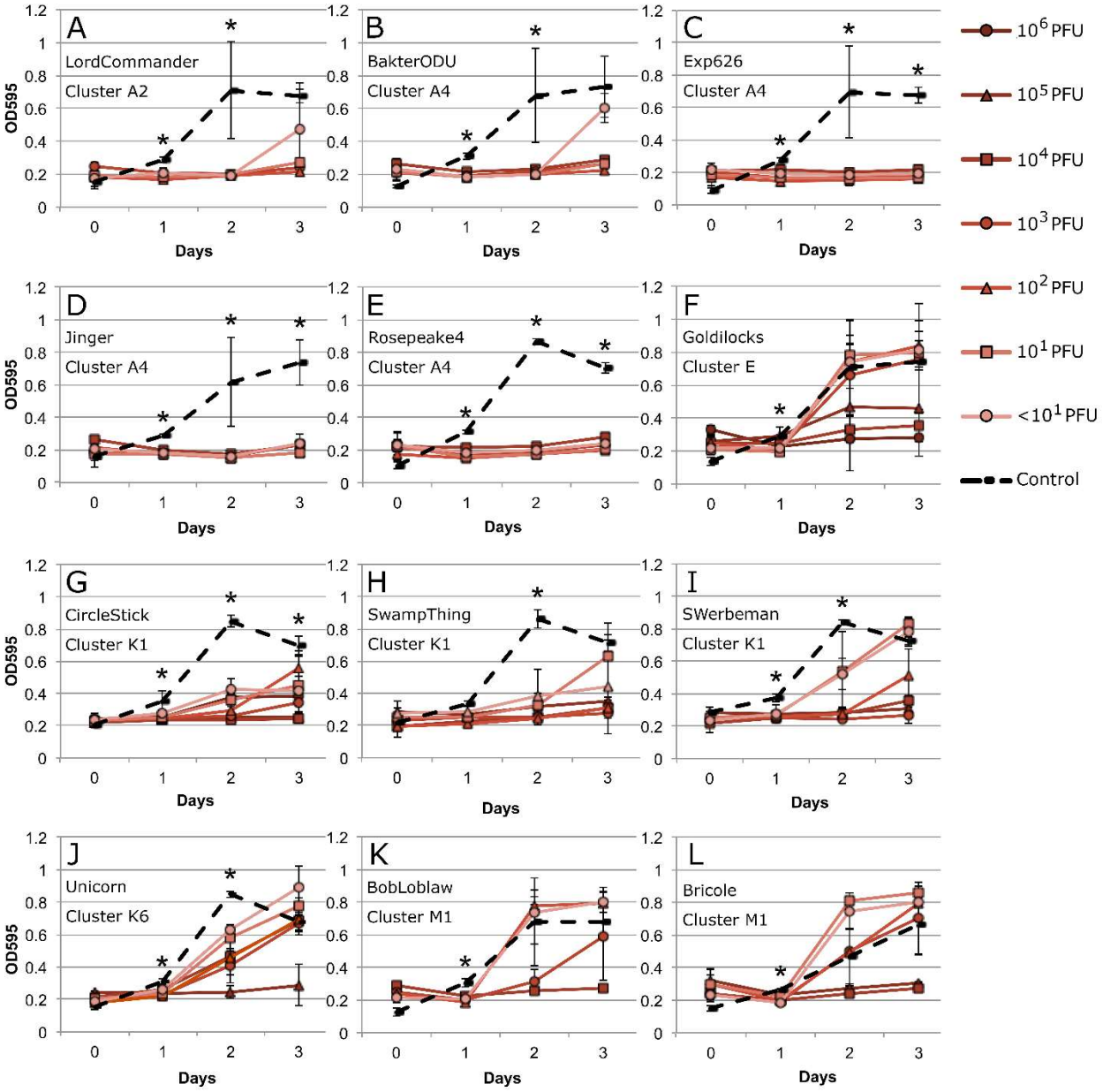
A
LordCommander
Cluster A2
B
BakterODU
Cluster A4
C
Exp626
Cluster A4
D
Jinger
Cluster A4
E
Rosepeake4
Cluster A4
F
Goldilocks
Cluster E
G
CircleStick
Cluster K1
H
SwampThing
Cluster K1
I
SWerbeman
Cluster K1
J
Unicorn
Cluster K6
K
BobLoblaw
Cluster M1
L
Bricole
Cluster M1
OD595
Days
10^6 PFU
10^5 PFU
10^4 PFU
10^3 PFU
10^2 PFU
10^1 PFU
<10^1 PFU
Control

FIG 4 Representative 96-well plate phage detection assays for each purified phage. The phages shown in each graph are (A) LordCommander (cluster A2), (B) BakterODU (cluster A4), (C) Exp626 (cluster A4), (D) Jinger (cluster A4), (E) Rosepeake4 (cluster A4), (F) Goldilocks (cluster E), (G) CircleStick (cluster K1), (H) SwampThing (cluster K1), (I) SWerbenman (cluster K1), (J): Unicorn (cluster K6), (K) BobLoblaw (cluster M1), (L) Bricole (cluster M1). The OD_{595} was recorded at the beginning of the experiment and at 24-hour intervals for four days; error bars represent +/- 1 standard deviation of triplicate wells. Dashed lines represent no-phage controls while solid lines represent experimental groups. Killing of bacteria by phages is inferred from OD measures that are significantly lower than the control values. Asterisks indicate days when all concentrations of phages (including $<10^1$ PFU) produced OD_{595} values significantly different from the control values.

Purification

Phages Exp626, Goldilocks, CircleStick, and Bricole (information can be found on phagesdb.org) were selected for the starting phage mixture and represent four different clusters (Table 5 and Fig. 5); clusters are groups of phages that share at least 50% sequence similarity (27). Plaque morphology of the individual phages used in the starting mixtures was determined based on visual examination after growth at 37°C for 24 hours on Luria agar plates (Fig. 5). Bricole produces very clear plaques with well-defined borders in a variety of sizes. CircleStick mostly produces very small, turbid plaques, although some are larger. Exp626 produces plaques of a wide range of sizes; most are small, but the occasional plaque is very large. The plaques are clear in the center, with a gradually increasing turbidity near the borders. Goldilocks produces plaques which sometimes have a bullseye appearance (a clear ring surrounding a turbid ring surrounding a clear center) in large and small sizes. The phage mixtures for Purifications A-C demonstrated highly variable plaque morphology prior to purification passages. Purification A solely produced clear plaques by round 3, suggesting isolation of Bricole; Purification B solely produced clear plaques by round 10, suggesting eventual isolation of Bricole; Purification C solely produced turbid plaques with clear centers by round 3, suggesting isolation of Exp626 (Fig. 6). Plaque size was not a good indication of phage isolation as the starting phages overlapped greatly in this characteristic.

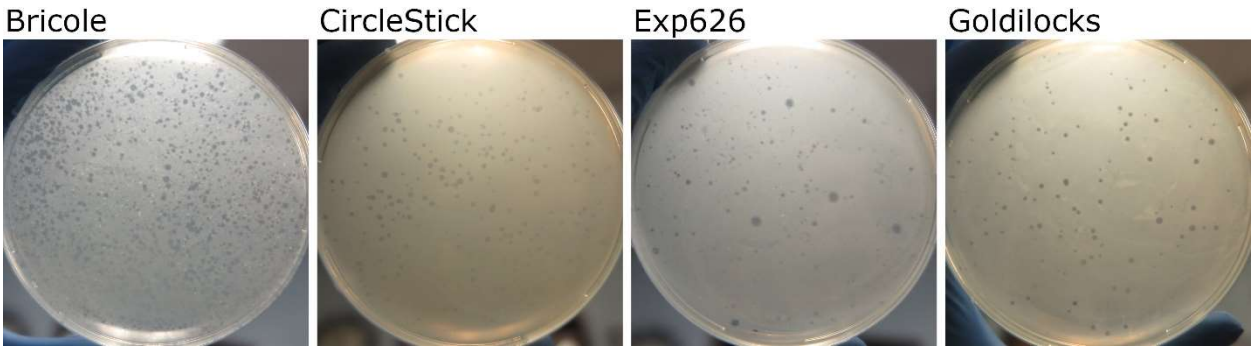

FIG 5 Plaque morphology of the four phages used in the purification study.

Because plaque morphology can be difficult to differentiate, restriction fragment length polymorphism (RFLP) and PCR were performed for a more detailed examination of phages present at each round of purification (Fig. 7 and 8). Both methods are dependent upon the DNA present in the sample and are extremely sensitive; however, this DNA may or may not be within an active phage particle.

For purification A, the RFLP indicated a mixture at rounds 0 and 1. Most of the other rounds only showed banding patterns consistent with Bricole; rounds 5 and 10 were unclear. The PCR analysis indicated a mixture at round 0 but only primers specific for Bricole produced bands in subsequent rounds. For purification B, the RFLP indicated a mixture at round 0, showed banding patterns consistent with Goldilocks at rounds 1 through 4, Bricole at rounds 6 through 9, and was unclear at rounds 5 and 10. The PCR analysis indicated a mixture at rounds 0 and 1, Goldilocks at rounds 2, 4, and 5, Bricole at rounds 8, 9, and 10, and a mixture of Goldilocks and Bricole at rounds 3 and 6. For purification C, the RFLP indicated a mixture at rounds 0, 1, and 2, had a banding

pattern consistent with Exp626 at rounds 3, 4, 7, 8, and 9, and was unclear at rounds 5, 6, and 10. The PCR analysis indicated mixtures of phages for the majority of rounds, but Exp626 only at rounds 7, 9, and 10.

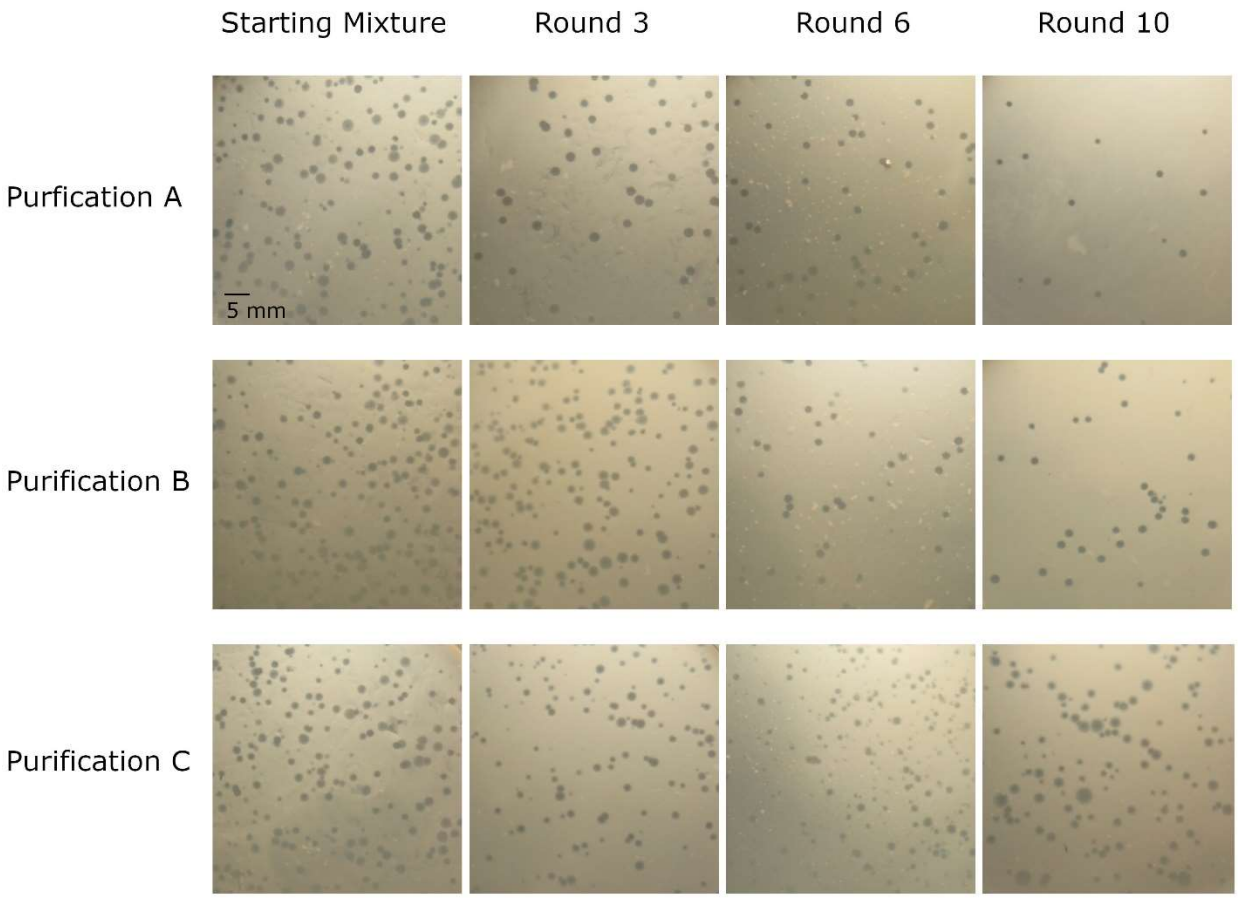

FIG 6 Plaque morphology for purifications A-C at each round of purification. The rows show purificatoins A, B, and C, while the columns show the starting mixture and rounds 3, 6, and 10. Plaque morphology is a commonly-used but inexact method of determining purification success.

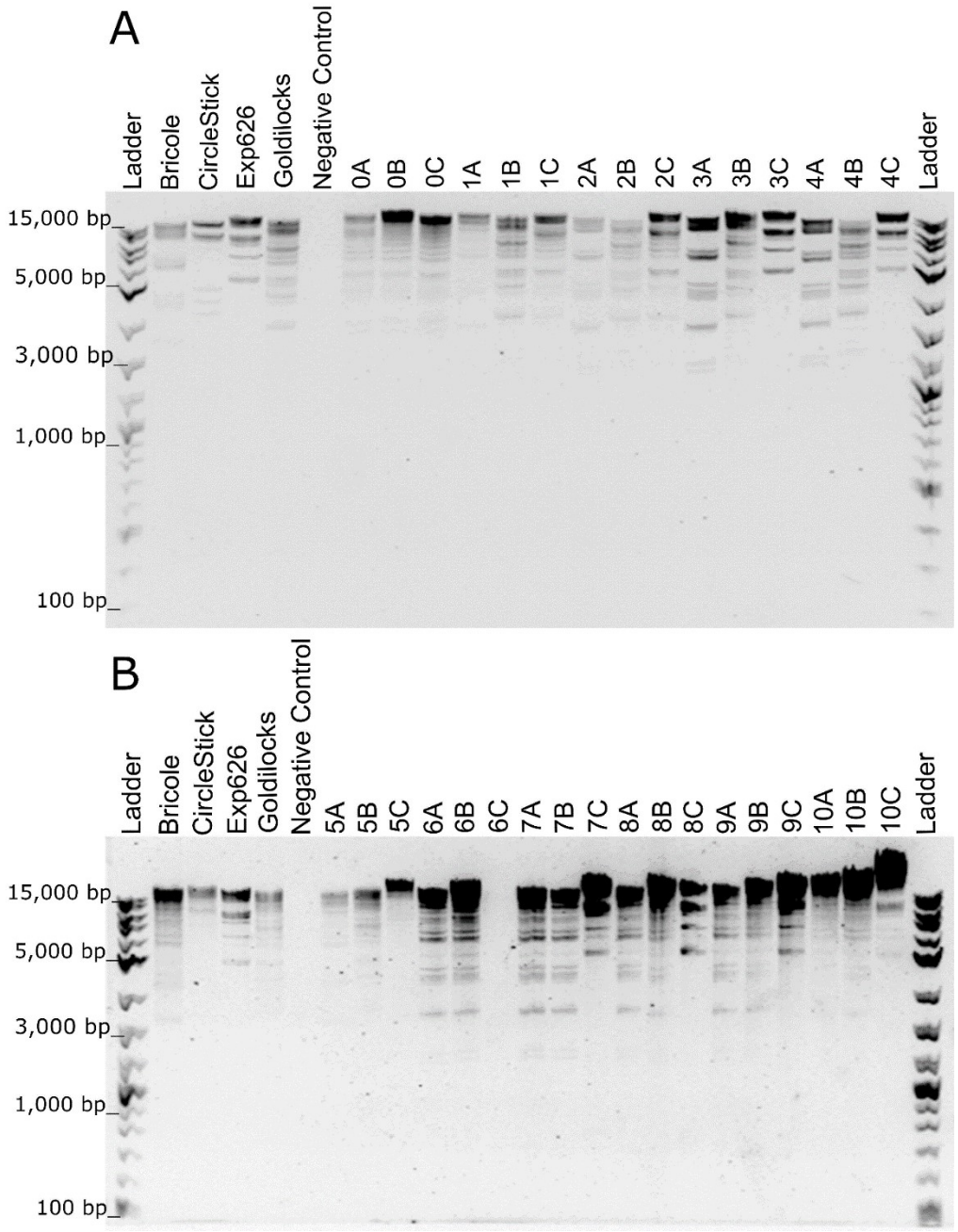

FIG 7 RFLP images for purifications A-C at each round of purification. Samples were digested with SexAI and EcoRI overnight at 37°C. In (A), the starting mixture of phages (round 0), and rounds 1 – 4 are shown. In (B), rounds 5 – 10 are shown. The ladder used was 1 Kb+ Plus DNA Ladder (Thermo Fisher).

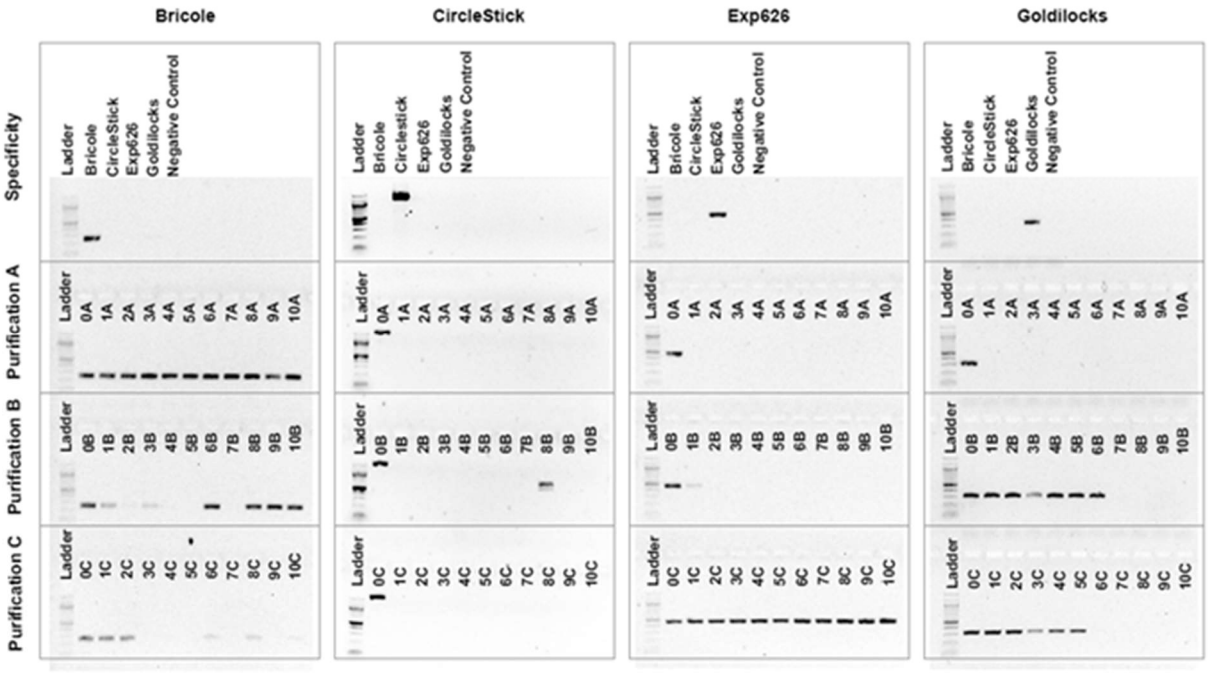

FIG 8 Agarose gel electrophoresis of PCR amplified products for purifications A-C at each round of purification. Large columns represent each primer set used (specific for Bricole, CircleStick, Exp626, or Goldilocks). The first row shows primer specificity and the following rows show the results for Purifications A-C. The controls in the top rows apply to the entire respective column for each primer set.

CircleStick-specific primers resulted in bands only for round 0 for all purifications (except for an unexpected band at Purification B round 8) but produced an unexpectedly large product size (band size is approximately 1,500 bp but was expected from *in silico* analysis to be 660 bp). Extensive troubleshooting did not resolve the issue, and it is possible the original CircleStick stock solution could have become contaminated by a different phage. It appears that whether CircleStick or another phage

was added at the beginning of the experiment, it was not detectable by the end of the experiment with the primers that were used, although it was detectable in initial phage mixtures.

3.4 DISCUSSION

The aim of this project was to discern whether the basic techniques of phage discovery can be performed using a liquid medium in multi-well plates. Eventually, a high-throughput pipeline could be developed for phage discovery in liquid medium in order to efficiently isolate phages, especially those that may not be as culturable with techniques requiring use of a solid medium.

In the multi-well plate enrichments, success in finding phage depended on the volume of environmental sample used, the source of the sample, and the starting concentration of phages in the sample (Table 8). For example, the soil of the Great Dismal Swamp generated more plaque-forming phage particles than the other sample sites. The Great Dismal Swamp is an area of low human influence, contains a great deal of decaying organic matter, and has a high moisture content; these factors may have influenced the number of phages present in the soil. The other sample sites (ODU and RES) also resulted in culturable phage; however, larger soil volumes appeared to be necessary for consistent phage isolation, likely because of lower starting phage concentrations in these samples. Although multi-well plates may be an effective option for many environmental samples, some samples with a low starting concentration of the desired phage may require higher sample volumes, which could preclude the use of multi-well plates.

The multi-well plate phage detection assay was highly sensitive; many of the phages required only a surprisingly small number of starting phages (<10) to produce reduction in optical density (Fig. 4). This level of sensitivity is not unprecedented; Rajnovic and colleagues found that as few as 10 phages per 220 µL could be detected by an optical density-based approach in 96-well plates, using the host *Escherichia coli* DSMZ 613 and the phage T4 (96). This experiment demonstrated that enrichments can be evaluated for their success or failure using a multi-well approach, which could also be used for other assays such as host range studies. Secondly, these experiments emphasize the diversity that exists within mycobacteriophages. Phages differed in the days at which host lysis could be measured and which phage concentrations produced measurable host lysis. It is currently unknown why these phages behaved differently in their ability to lyse the solution of host bacteria, but it is likely due to factors such as adsorption rate, growth rate, burst size, and ability to penetrate biofilms (57). Although the lytic phages (BobLoblaw and Bricole) produced the larger average plaque sizes (Fig. 5) than the temperate phages (LordCommander, BakterODU, Exp626, Jinger, Rosepeake4, Goldilocks, CircleStick, SwampThing, SWerbenman, and Unicorn), they did not did not produce a more pronounced difference between experimental and control groups in the 96-well plate detection assays (Fig. 4K-L). Host lysis on solid and liquid media may be affected by more factors than are currently known. Phages of the same cluster tended to share characteristics such as phage concentration required for detection, number of days the difference between experimental and control groups was maintained, and how substantial this difference was. For example, for the cluster A phages used (LordCommander, BakterODU, Exp626, Jinger, and Rosepeake4), the

experimental and control group absorbance values tended to diverge one day after inoculation and remain distinct though the remainder of the experiment, even for very low phage concentrations. Because clusters are based on genetic relatedness, it can be inferred that many characteristics related to phage growth and host-pathogen interactions are influenced by the genetic makeup of the phage.

The 96-well plate purification assay was the most novel of the procedures described. It was surprising how quickly individual phages could be purified out of a phage mixture (all three purifications appeared to reach individual phage isolation within 10 rounds of purification); especially interesting was the speed with which Bricole was isolated in purification A (two rounds of purification) (Fig. 5-8). Phage characteristics such as burst size, lifestyle, and growth rate might play a large role in determining which phage is purified and how many passages are required for purification; in addition, different environmental conditions may favor different phages (57). It was interesting that in purification B, Goldilocks seemed to have outcompeted the other phages until round 6, at which point Bricole rapidly became the predominant phage. This implies that, even when one is confident that a phage solution is pure, minute numbers of contaminating phage could suddenly gain an advantage and overwhelm the original predominant phage. By performing excessive rounds of purification, one may inadvertently influence which phage becomes isolated, possibly hindering the ability to isolate phages that have a competitive disadvantage under laboratory conditions.

Neither RFLP nor PCR were wholly consistent in their ability to discern which phages were present in the solution (Fig. 7-8). Only a small portion of the original mixture is used in both techniques; if only a few phage particles are present in a

solution, these may or may not be transferred to the RFLP or PCR reaction. Since phages can propagate exponentially, a small number of phages in a solution may be undetectable by current methods but able to propagate to larger numbers over time. However, RFLP and PCR are likely more accurate than the typical methods for determining phage purification, such as plaque morphology and electron microscopy.

The benefits of using multi-well plates include greater efficiency, reduced cost, ability to automate procedures, increased number of samples that can be screened, and the possibility of isolating a greater diversity of phages. It will be valuable to convert as many phage procedures as possible to an efficient, multi-well format in anticipation of increased demand for phage isolation. It is also important to have multiple possible procedures that achieve the same goal in different ways, as some phages may be more amenable to certain conditions or may be fastidious in their requirements. A combination of efficiency and flexibility will allow rapid, practical, and effective isolation of phages for phage therapy, biocontrol, and other applications.

CHAPTER 4

PHYLOGENY AND EVOLUTION OF PROPHAGES IN 49 GENOMES OF *MYCOBACTERIUM MARINUM*

4.1 INTRODUCTION

Phages are viral pathogens of bacteria. In the lytic cycle, the host cell's molecular machinery produces progeny phage, which exit by lysing the bacterial cell, while in the lysogenic cycle, the phage incorporates its genetic sequence into the host DNA where the sequence (now called a prophage) is replicated along with the host's genetic material until induction occurs (68). Prophages can be induced by various environmental stressors or spontaneous induction can occur (16). Prophages are often strain-specific and account for a large percentage of the genetic variability between bacterial strains (97). Prophages may provide advantages to the host, such as encoding virulence factors and superinfection exclusion systems, and/or may have disadvantages such as encumbering the cell with extra DNA or disrupting a necessary bacterial gene (98). Prophages may degrade via the accumulation of point mutations, insertions, deletions, and rearrangements if there is not adequate selective pressure to maintain prophage genes (98).

Unfortunately, there is no definitive criteria to identify prophages at the genetic level, and current prophage-detecting software is useful but imperfect (99). Certain prophage genes are well conserved, such as the large terminase subunit gene, the portal protein gene, the head maturation protease gene, the coat protein gene, tail shaft genes, the tape measure gene, and tail fiber genes; however, most prophage genes are

not universally present and many have bacterial homologues (100). For example, the integrase gene is similar to non-phage elements such as transposases (101). Still, the most accurate and widely-used criterion for phage identification is similarity of the sequence to sequences of known phages and prophages; of course, this neglects an unknown number of prophages due to the limited number of identified and sequenced phages (99).

Few systematic analyses of prophages have been undertaken; prophage analysis often entails examining prophage number, similarity, insertion site, attachment (*att*) sequences, and mosaicism, but few seek to determine the number of integration events or compare host and prophage phylogeny in a large number of genomes. Bobay and colleagues identified over 300 prophages in 85 genomes of *Enterobacteriaceae* and determined that these prophages exhibit strong sequence conservation, a bimodal size distribution, evidence of purifying selection on core phage genes, and a likely vertical inheritance (102). Shaaban and colleagues, who analyzed 14 *Escherichia coli* O157 prophages (103), and Fu and colleagues, who analyzed 171 *Streptococcus mutans* prophages (104), each performed a phylogenetic analysis and found clustering based on sequence similarity of the prophages. Casjens and Grose identified 9,371 prophages in 3,298 *Salmonella* genomes; they determined that similar prophages (P2 and P22) are found in diverse hosts and analyzed the phylogenies of individual prophage gene types (105). Vale and colleagues found 28 prophages in 28 *Helicobacter pylori* genomes, compared the insertion sites, and performed a phylogenetic analysis that demonstrated that prophage similarity in these genomes was correlated with the geographic area in which the host was collected (106). These

investigations are noteworthy because they include in-depth analyses of prophage phylogenetics using large numbers of host genomes.

The genus *Mycobacterium* includes the pathogens that cause tuberculosis (*M. tuberculosis*/*M. bovis*) and leprosy (*M. leprae*), but many aquatic mycobacteria, such as *M. marinum*, are also significant pathogens, affecting both aquatic animals and humans (1, 95). Nontuberculous mycobacterial infections are becoming increasingly prevalent in humans (44). Mycobacteria are resistant to many stressors including several antibiotics due to their complex cell wall, making treatment difficult (46). Phage therapy and treatment with phage components have been proposed as solutions to mycobacterial infections in both humans and nonhuman animals (47, 49).

Despite the medical, ecological, and commercial importance of this genus, there have been few large-scale analyses of prophage phylogeny in mycobacteria. Hendrix and colleagues compared phage and prophage sequences from *Mycobacterium*, *Haemophilus*, *Streptomyces*, and *Escherichia*, and found mosaicism from a common gene pool due to horizontal gene transfer, which occurs both between phages and between a phage and its host (107). In 2014, Fan and colleagues analyzed 33 prophages from 30 genomes of 23 species of mycobacteria (including saprophytic species, animal pathogens, and human pathogens); 11 were considered intact prophages (108). Comparing a phylogenetic tree of the prophage integrase genes to a tree of the bacterial genomes, the authors concluded that prophage similarity was not a good predictor of overall genome similarity (108). In 2016, Fan and colleagues compared two prophages across *M. tuberculosis* complex strains, and found insertions, deletions, and reorganization in a manner independent of the relatedness of the hosts

(109). Sassi and colleagues analyzed the prophages of 48 isolates of *M. abscessus* (which includes two subspecies) (110). Based on full protein sequence comparison, they found a correlation between prophage similarity and subspecies of *M. abscessus*; many horizontal gene transfer events were also inferred (110). Voronina and colleagues found that prophage profile was associated with sub-strain in *M. bovis* (111). Bouam and colleagues compared three strains of *M. simiae* complex and found extensive differences in the prophages found in each strain (112).

In this study, prophages were identified in 49 sequenced genomes of *Mycobacterium ulcerans-Mycobacterum marinum* complex bacteria (MuMC). The prophages were analyzed for sequence length, gene content, insertion site, and other characteristics. Phylogenetic networks of the prophages and their hosts were constructed to determine the relatedness of the prophages and the relative age of the prophage integration events relative to the diversification of the MuMC strains.

4.2 MATERIALS AND METHODS

DNA Extraction

Cultured mycobacteria were grown with previously published techniques (113). Briefly, mycobacteria were grown in Middlebrook 7H9 broth with OADC enrichment and 0.1% Tween-20 to turbidity with shaking at 24°C. Bacteria were pelleted by centrifugation and washed in Butterfield's Phosphate Buffered Saline + 0.1% Tween-20 (BPBST). Bacterial pellets were extracted for DNA using the DNeasy extraction kit (QIAGEN, Valencia, CA) preceded by beadmilling with 0.1mm Si-Zi beads for 40s at

5000 rpm. DNA quantity and quality were determined using NanoDrop (ThermoFisher, Waltham, MA).

DNA Sequencing

DNA sequencing was performed by Dr. David Gauthier at Old Dominion University, Norfolk, Virginia. Isolates were sequenced to at least 50× coverage with either Illumina HiSeq (2×150bp) or MiSeq (2×250bp) platforms. Reads were assembled with Celera WGS (v8.1) (114). In genomes assembled to draft contig level, synteny with closely-related strains was used to infer the contig arrangement. Sequence data for certain isolates (Table 9) were obtained from USDA Animal Plant Health Inspection Service-National Veterinary Services Laboratory-Diagnostic Bacteriology Laboratory (USDA-NVSL-DBL) and are deposited in the NCBI Sequence Read Archive (SRA). Previously deposited genomes were also used in the analysis, as indicated in Table 9. Sequence assemblies were annotated with Prokka (115). Genome sequences for *M. shottsii* and *M. marinum* ATCC927 were generated using PacBio sequencing. Geneious version 9.1.8 (Biomatters Ltd., Auckland, New Zealand) was used for manual curation, alignments, and basic sequence statistics (Table 9) (116).

TABLE 9 MuMC strain characteristics, including source, length, number of contigs, number of prophages, and number of putative CRISPR sequences. Number of prophages was determined by manual curation of PHASTER results. Number of CRISPR sequences was determined by level 3 and 4 results from CRISPRFinder. SB = striped bass (*Morone saxatilis*), HSB = hybrid striped bass (*Morone saxatilis* x *Morone chrysops*), USDA = USDA Animal Plant Health Inspection Service-National Veterinary Services Laboratory-Diagnostic Bacteriology Laboratory (USDA-NVSL-DBL), ATCC = American Type Culture Collection (ATCC), OST = Ostland and colleagues, 2008, VDH = Virginia Department of Health, MSU = Dr. Frank Austin at Mississipi State University, SHEDD = Shedd Aquarium (Chicago, IL), VIMS = Virginia Institute of Marine Science

Strain name	Source organism	Isolation source	Accession /ID	Length (bp)	Contigs	Prophages	CRISPR
12-0014	Python	USDA	SRS1039989	6,199,210	557	5	0
12-0017	Frog	USDA	SRS1039988	6,183,441	760	5	0
12-1908	Toad	USDA	SRS1039986	6,254,872	469	4	0
12-1909	Toad	USDA	SRS1039985	6,267,052	436	4	0
12-8618	Frog	USDA	SRS1039983	6,155,244	250	3	0
12-9333	Frog	USDA	N/A	5,358,708	1429	3	0
13-0905	Toad	USDA	SRS1039982	6,225,813	406	4	0
13-3845	Toad	USDA	SRS1039981	6,137,551	245	3	0
13-5434	Frog	USDA	SRS1039980	5,975,363	488	2	0
14-0534	Kangaroo	USDA	SRS1039979	6,289,537	310	4	0
1218R	Fish		This study	6,077,918	370	3	0
ATCC15069	Human	ATCC	This study	6,400,915	170	3	0
ATCC25039 (H1726)	Armadillo	ATCC	This study	6,452,623	108	4	0
ATCC33728 (*M. ulcerans*)	Human	ATCC	This study	5,841,623	212	1	0
ATCC927	Fish	ATCC	This study	6,674,882	3	5	0
ATCC11564	Human	ATCC	This study	6,333,791	331	3	0

TABLE 9 continued

Strain name	Source organism	Isolation source	Accession /ID	Length (bp)	Contigs	Prophages	CRISPR
ATCC29254	Human	ATCC	This study	5,839,079	353	1	0
C7	SB	VIMS	This study	6,133,824	251	4	0
C35	SB	VIMS	This study	6,184,366	62	5	1
Europe	Fish	VIMS	GCA_000 419315.1	6,029,340	4	2	0
KST94	Fish	OST	This study	6,108,059	187	1	0
KST214	HSB	OST	This study	5,923,318	269	1	0
KST266	Fish	OST	This study	5,957,871	276	1	0
KST417	Fish	OST	This study	6,025,757	259	1	0
KST687	Fish	OST	This study	6,096,129	219	1	0
L50	SB	VIMS	This study	6,156,803	179	1	0
M	Human	VIMS	NC_0106 12	6,636,827	1	4	0
M2	Fish	SHEDD	This study	6,223,536	93	2	0
M4	Fish	SHEDD	This study	5,498,189	163	2	0
M11 (4.6)	Human	VDH	This study	6,159,885	236	2	0
M12 (MSU90)	Fish	MSU	This study	6,296,274	253	2	0
M13 (MSU97)	Fish	MSU	This study	6,359,613	207	4	0
M30 (VIMS9)	SB	VIMS	This study	6,366,776	276	5	2
M50	SB	VIMS	This study	6,368,089	255	5	2
M59	SB	VIMS	This study	6,285,761	108	2	0
M324-958	SB	VIMS	This study	6,467,262	320	3	2
M453-022	SB	VIMS	This study	6,161,761	309	1	0
MB2	Fish	VIMS	GCA_000 419335.1	6,134,390	3	2	0
M. liflandii (Mli)	Frog	VIMS	NC_0201 33	6,208,955	1	2	0
M. pseudoshottsii (Mps)	SB	VIMS	This study	6,096,009	260	2	0
Mycobacterium sp. 012931 (Msp)	Fish	VIMS	GCA_000 419295.1	5,759,365	1	2	0
MSS4	Fish	OST	This study	6,132,847	295	2	0
M. ulcerans (Mul)	Human	VIMS	NC_0059 16	5,631,606	1	1	0
R106	SB	VIMS	This study	6,585,954	320	4	3
R171	SB	VIMS	This study	6,116,290	331	2	0
Rp72a	SB	VIMS	This study	6,137,331	308	2	0
M. shottsii (TM48)	SB	VIMS	This study	5,956,408	1	3	0

Criteria for Prophage Identification

The program PHASTER (PHAge Search Tool - Enhanced Release) identified possible prophages (117). Added to this preliminary list were the prophages identified in *M. marinum* M reported by Stinear and colleagues (118). Attachment (*att*) sites are the sequences on either side of the prophage that represent the location of prophage insertion; PHASTER was able to determine the putative *att* sites for some of the prophages. The progressiveMauve algorithm in Mauve 2.4.0 was used to generate alignments of the bacterial genomes so that the presence of inserted sequences in the putative prophage insertion sites could be identified in the other genomes (119). These sequences represent potential prophages too highly degraded to be identified by PHASTER.

The list of putative sequences was curated using NCBI Protein BLAST (120, 26) (a search of genomes in the NCBI database) and the Actinobacteriophage Database Protein BLAST (a search of mycobacteriophage genomes) (92). Casjens describes certain genes as “cornerstone” gene, i.e. genes that are more highly conserved in phages and less likely to have bacterial homologues; these genes encode the scaffolding protein, terminase, portal protein, head/capsid maturation protease, coat protein, tapemeasure, and structural proteins of the capsid, tail, and tail fibers (100). Integrase and lysin genes are considered supporting, but not sufficient, evidence that a sequence is a prophage due to bacterial homologues (100). Integrases are especially difficult to distinguish due to their similarity with transposases and other DNA-excision genes (101).

In order to be included in this study, the sequence was required to fulfill at least two of the following three criteria:

1. Considered a prophage by PHASTER or by previously published annotation, e.g. Stinear and colleagues (118)
2. At least one match to a gene found in a cultured phage (e-value < 1×10^{-4}, bit score > 50)
3. Shares an insertion site with a prophage that has at least one match to a cornerstone gene (listed above) (e-value < 1×10^{-4}, bit score > 50)

Prophages were named with the name of the bacterial strain followed by the letter p and a number representing the order of discovery. Insertion sites were named with capital letters in the order of discovery. A prophage was considered possibly intact if it was at least 20 Kb in nucleotide length and contained genes from at least 5 of the following classes: scaffolding, terminase, portal, head/capsid maturation protease, coat protein, tapemeasure, capsid, tail, tail fiber, integrase, and lysin.

Phylogenetic Networks of Bacterial and Prophage Genomes

The progressiveMauve algorithm in Mauve 2.4.0 was used to generate alignments of the bacterial and prophage genomes (119). Distance matrices from each alignment were used to generate a splits graph under the NeighborNet algorithm in SplitsTree (121). Split-decomposition phylogenetic networks are useful for inferring complex evolutionary processes such as horizontal gene transfer that are not compatible with traditional bifurcating phylogenetic trees (122).

Prophage Gene Analyses

The Actinobacteriophage Database protein BLAST (phagesdb.org) was used to find the closest match to each prophage gene (e-value < 1×10^{-4}, bit score > 50). Geneious version 9.1.8 (Biomatters Ltd., Auckland, New Zealand) (117) was used to construct gene maps for a representative prophage of insertion sites A (ATCC927_p2), B (ATCC11564_p2), and C (ATCC927_p1).

Phylogenetic trees were created for the following prophage gene types: integrase, tapemeasure, terminase (large subunit), and portal. These trees were compared to the network of full prophage sequences, to establish whether the prophages have a mosaic structure and whether the diversity of any of these genes recapitulates the full sequence diversity of their respective prophages. Not every gene assigned to these classifications was used; only sequences with at least 50% identity to each other were used to avoid comparing nonhomologous sequences. The integrases were split into two groups based on sequence identity, using MUSCLE (123) in Geneious version 9.1.8. PhyML (124) was used to perform maximum-likelihood analysis subsequent to determination of the optimum evolutionary model with JModelTest (125).

On the Actinobacteriophage Database website (phagesdb.org), a protein BLAST search identified close matches to each prophage gene and the phage and cluster to which each gene match belonged (e-value < 1e-4, bit score > 50). Clusters are groups of phages that have a sequence similarity of at least 50% (27). A pie chart of gene matches was made in Microsoft Excel 14.1.0 for a representative prophage from each of insertion sites A (ATCC927_p2), B (ATCC11564_p2), and C (ATCC927_p1).

CRISPR Identification

The program CRISPRFinder located putative CRISPR sequences in some of the bacterial genomes (126). Only sequences designated level 3 or 4 (the highest levels of certainty) were included in the analysis. The number of CRISPR sequences identified in each genome was compared to the number of prophage sequences found in those genomes.

Microsoft Excel 14.1.0 was used to generate the scatterplot, heatmaps, and pie charts, and to calculate the Spearman's rank correlation coefficient for the scatterplot.

4.3 RESULTS

The MuMC genomes contained between one and five prophages, which were found in nine different insertion sites (identified as A through I), relative to the representative genome *M. marinum* M (Fig. 9). A total of 134 presumptive prophage sequences were identified, only 9 of which are possibly intact based on their sequence size and types of phage genes present (Table 10).

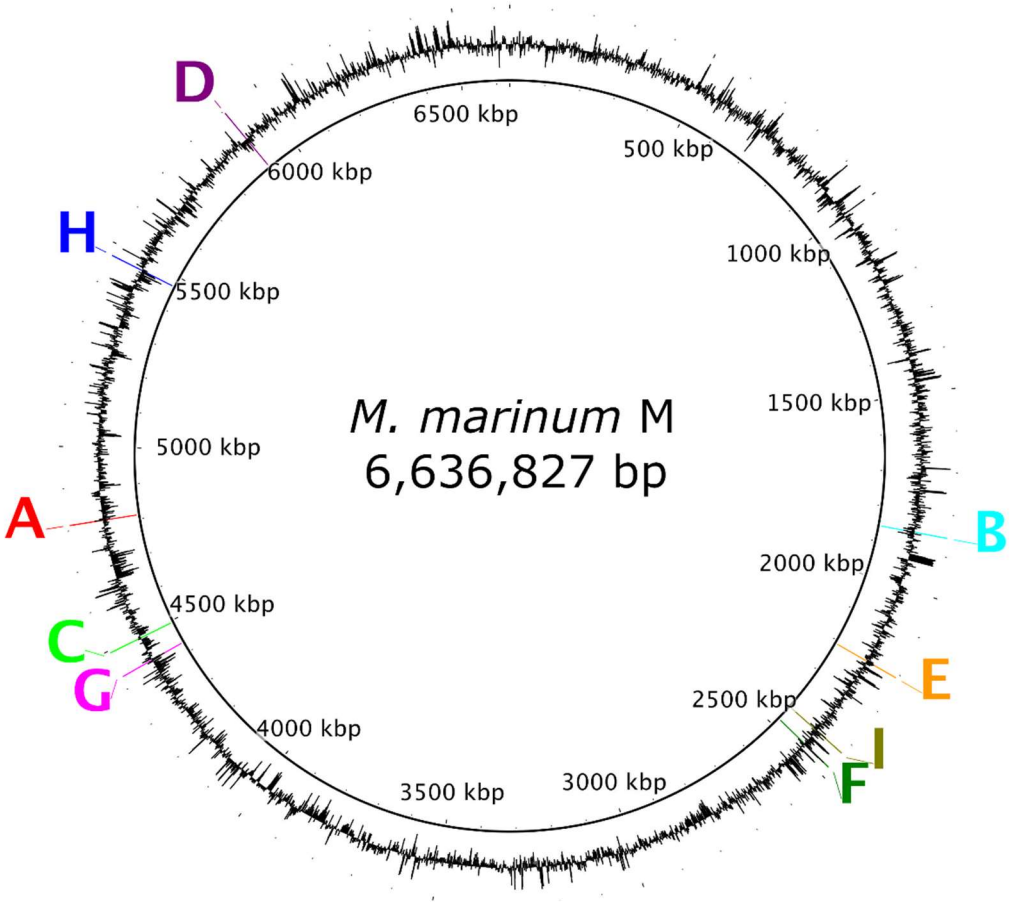

FIG 9 Circular genome of *M. marinum* M showing prophage insertion sites (BRIG). Not all insertion sites are occupied by prophages in *M. marinum* M but are shown here for reference.

TABLE 10 Prophage characteristics. Information was derived using Geneious version 9.1.8 (Biomatters Ltd., Auckland, New Zealand). Intactness was defined according to nucleotide length and types of genes in the prophage, as described in Materials and Methods

Prophage name	%GC	Insertion site	Length (bp)	Number of CDS	Possibly intact?
ATCC927_p2	63.6%	A	58737	88	Yes
ATCC25039_p1	63.6%	A	58602	88	Yes
C35_p1	63.4%	A	57797	91	Yes
Davis_p1	63.7%	A	57395	85	Yes
13-0905_p1	63.6%	A	57051	80	Yes
M13_p1	63.5%	A	56937	87	Yes
M_p1*	63.5%	A	56937	88	Yes
12-0014_p2	62.5%	B	8049	11	No
12-0017_p2	63.4%	B	10467	13	No
12-1908_p2	66.5%	B	5242	8	No
12-1909_p2	66.5%	B	5242	8	No
1218R_p1	63.5%	B	57528	88	Yes
14-0534_p1	62.9%	B	9301	11	No
ATCC11564_p2	63.5%	B	57643	90	Yes
C35_p2	64.0%	B	11492	9	No
C7_p1	64.0%	B	11492	9	No
M30_p1	63.3%	B	1368	1	No
M324-958_p1	64.1%	B	2952	3	No
M50_p1	63.3%	B	1368	1	No
R106_p1	63.3%	B	1368	1	No
ATCC927_p1	63.2%	C	50594	77	Yes
12-0014_p3	62.7%	D	19224	22	No
12-0017_p3	62.7%	D	19405	22	No
12-1908_p1	60.9%	D	36424	30	No
12-1909_p1	60.9%	D	36618	30	No
13-0905_p2	61.3%	D	19311	21	No
13-5434_p2	62.9%	D	11906	14	No
14-0534_p2	63.0%	D	19669	23	No
ATCC927_p3	61.2%	D	18602	19	No
ATCC11564_p1	62.2%	D	19370	18	No
ATCC15069_p3	62.0%	D	8001	7	No
C7_p2	62.0%	D	8001	7	No
C35_p3	62.0%	D	8001	7	No
Europe_p1	61.1%	D	18380	18	No
M13_p2	62.6%	D	18509	21	No
M30_p2	68.8%	D	38746	44	No
M50_p2	68.8%	D	38781	44	No
M_p3*	62.7%	D	18992	22	No
Mli_p1	62.7%	D	19786	21	No
Mps_p1	62.8%	D	18405	20	No

TABLE 10 continued

Prophage name	%GC	Insertion site	Length (bp)	Number of CDS	Possibly intact?
Msp_p1	62.8%	D	18396	20	No
TM48_p2	59.8%	D	16894	17	No
12-0014_p1	65.1%	E	12667	18	No
12-0017_p1	64.6%	E	11589	20	No
12-0014_p4	58.0%	F	3068	3	No
12-0017_p4	58.1%	F	3072	3	No
12-1908_p3	62.3%	F	13856	25	No
12-1909_p3	62.4%	F	13066	24	No
12-8618_p1	57.4%	F	8691	8	No
12-9333_p2	57.4%	F	8691	8	No
13-0905_p3	62.3%	F	13231	24	No
13-3845_p2	58.0%	F	9661	9	No
13-5434_p1	62.5%	F	6378	10	No
14-0534_p3	60.0%	F	4258	4	No
ATCC15069_p2	61.8%	F	8914	15	No
C7_p3	61.6%	F	9322	14	No
C35_p4	61.6%	F	9322	14	No
Davis_p2	65.6%	F	10165	16	No
E11_p1	62.9%	F	15346	18	No
KST94_p1	65.6%	F	9429	18	No
KST214_p1	65.6%	F	9429	18	No
KST266_p1	65.6%	F	9429	18	No
KST417_p1	65.6%	F	9429	18	No
KST687_p1	65.6%	F	9429	18	No
L50_p1	65.6%	F	9429	18	No
M2_p1	57.1%	F	5828	4	No
M4_p1	65.6%	F	9419	17	No
M11_p1	57.7%	F	6699	5	No
M12_p1	61.8%	F	9973	15	No
M30_p3	59.9%	F	4092	5	No
M50_p3	58.7%	F	3376	4	No
M59_p1	62.1%	F	10890	13	No
M324-958_p2	58.6%	F	3064	3	No
M453-022_p1	65.6%	F	9429	18	No
MB2_p1	57.1%	F	5828	4	No
MSS4_p1	61.5%	F	9071	14	No
R106_p3	60.0%	F	4127	5	No
R171_p1	65.6%	F	9429	18	No
Rp72a_p1	65.6%	F	9429	18	No
TM48_p3	63.3%	F	9490	12	No
12-0014_p5	61.8%	G	14569	17	No
12-0017_p5	61.8%	G	14569	16	No
12-1908_p4	61.2%	G	13440	12	No
12-8618_p3	57.8%	G	20385	23	No
12-9333_p3	55.3%	G	9149	10	No
13-0905_p4	61.0%	G	7616	9	No
13-3845_p3	57.3%	G	16468	19	No
14-0534_p4	61.8%	G	14569	16	No

TABLE 10 continued

Prophage name	%GC	Insertion site	Length (bp)	Number of CDS	Possibly intact?
1218R_p2	60.0%	G	12257	13	No
ATCC927_p4	60.7%	G	11014	10	No
ATCC11564_p3	59.8%	G	11807	13	No
ATCC15069_p4	61.7%	G	6288	8	No
ATCC29254_p1	59.2%	G	4830	9	No
ATCC33728_p1	63.6%	G	429	1	No
C7_p4	63.6%	G	654	2	No
C35_p5	63.6%	G	654	2	No
Davis_p3	62.7%	G	579	2	No
E11_p2	61.2%	G	9053	8	No
Europe_p2	63.6%	G	654	2	No
M2_p2	63.4%	G	579	2	No
M4_p2	52.5%	G	1207	3	No
M11_p2	63.4%	G	579	2	No
M12_p2	56.5%	G	3496	5	No
M13_p3	61.7%	G	25681	25	No
M30_p4	62.8%	G	5290	3	No
M50_p4	62.8%	G	5290	3	No
M59_p2	63.2%	G	579	2	No
M324-958_p3	68.5%	G	2418	2	No
M_p4*	61.7%	G	25681	25	No
MB2_p2	63.4%	G	579	2	No
Mli_p2	63.8%	G	654	2	No
Msp_p2	63.6%	G	654	2	No
MSS4_p2	56.5%	G	3496	5	No
Mul_p1	64.1%	G	429	1	No
R106_p4	68.5%	G	2418	2	No
R171_p2	52.5%	G	1207	3	No
Rp72a_p2	52.5%	G	1207	3	No
TM48_p1	65.1%	G	7244	8	No
12-8618_p2	58.7%	H	4280	6	No
12-9333_p1	58.7%	H	4280	6	No
13-3845_p1	58.7%	H	4279	7	No
ATCC927_p5	58.5%	H	3201	3	No
ATCC25039_p4	58.5%	H	3201	3	No
M13_p4	58.9%	H	3672	4	No
M30_p5	57.9%	H	7023	8	No
M50_p5	57.9%	H	7023	8	No
M_p2*	61.3%	H	10348	13	No
ATCC15069_p1	63.2%	I	15517	28	No

*Phages marked with an asterisk are described by Stinear and colleagues (118). The prophage M_p1 is equivalent to phiMmar02; M_p2 is equivalent to phiMmar03; M_p3 is equivalent to phiMmar05; M_p4 is equivalent to phiMmar10.

The most intact prophages (based on number and types of genes present) were located in insertion sites A, B, and C. The remaining prophages appeared to be in varying stages of degradation; some were as small as one integrase gene and identified only because of a shared insertion site. Color-coded gene maps were created to emphasize the synteny (conservation of gene order) between prophage types (Fig. 10). Synteny was observed in the possibly intact prophages of insertion sites A, B, and C. One representative prophage from each of those insertion sites was used to make a gene map, colored according to gene type. Most of the gene types aligned very well along the three prophages, except for the reversed placement of the integrase in insertion site C.

The prophages were highly mosaic in their genetic structure when compared to mycobacteriophages in the Actinobacteriophage Database (phagesdb.org) (Fig. 11). Many of the prophage genes shared high amino acid sequence similarity to phage genes in that database; however, none of the prophages as a whole were a close enough match (>50% sequence similarity) to any phage cluster to designate a cluster assignment (27).

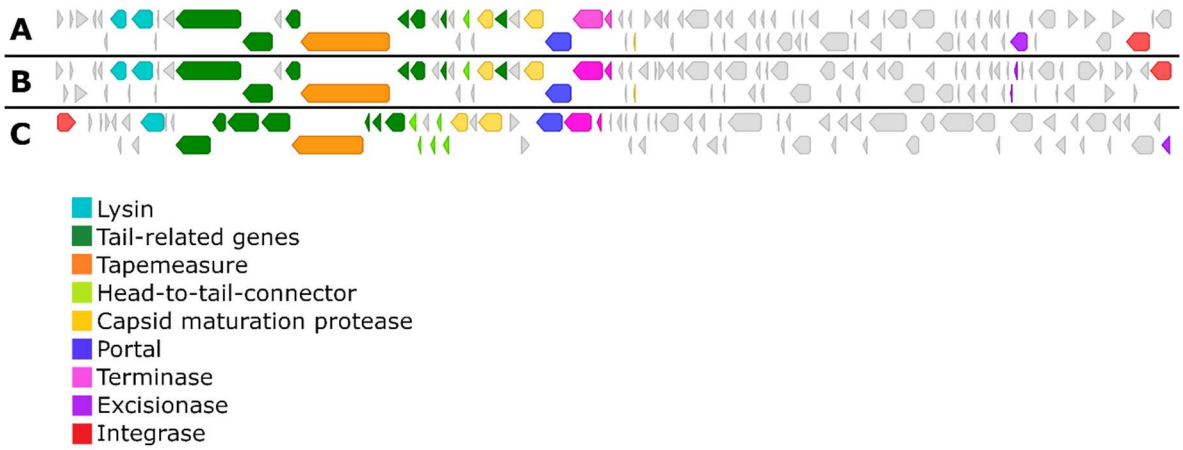

FIG 10 Map of tapemeasure, integrase, lysin, terminase and portal, and capsid- and tail-related genes for a representative prophage from insertion sites A (ATCC927_p2), B (ATCC11564_p2), and C (ATCC927_p1). Annotations are based on phagesdb.com protein BLAST; e value < 1×10^{-4}, bit score > 50. Multiple sequence alignment was performed using MUSCLE in Geneious version 9.1.8 (Biomatters Ltd., Auckland, New Zealand).

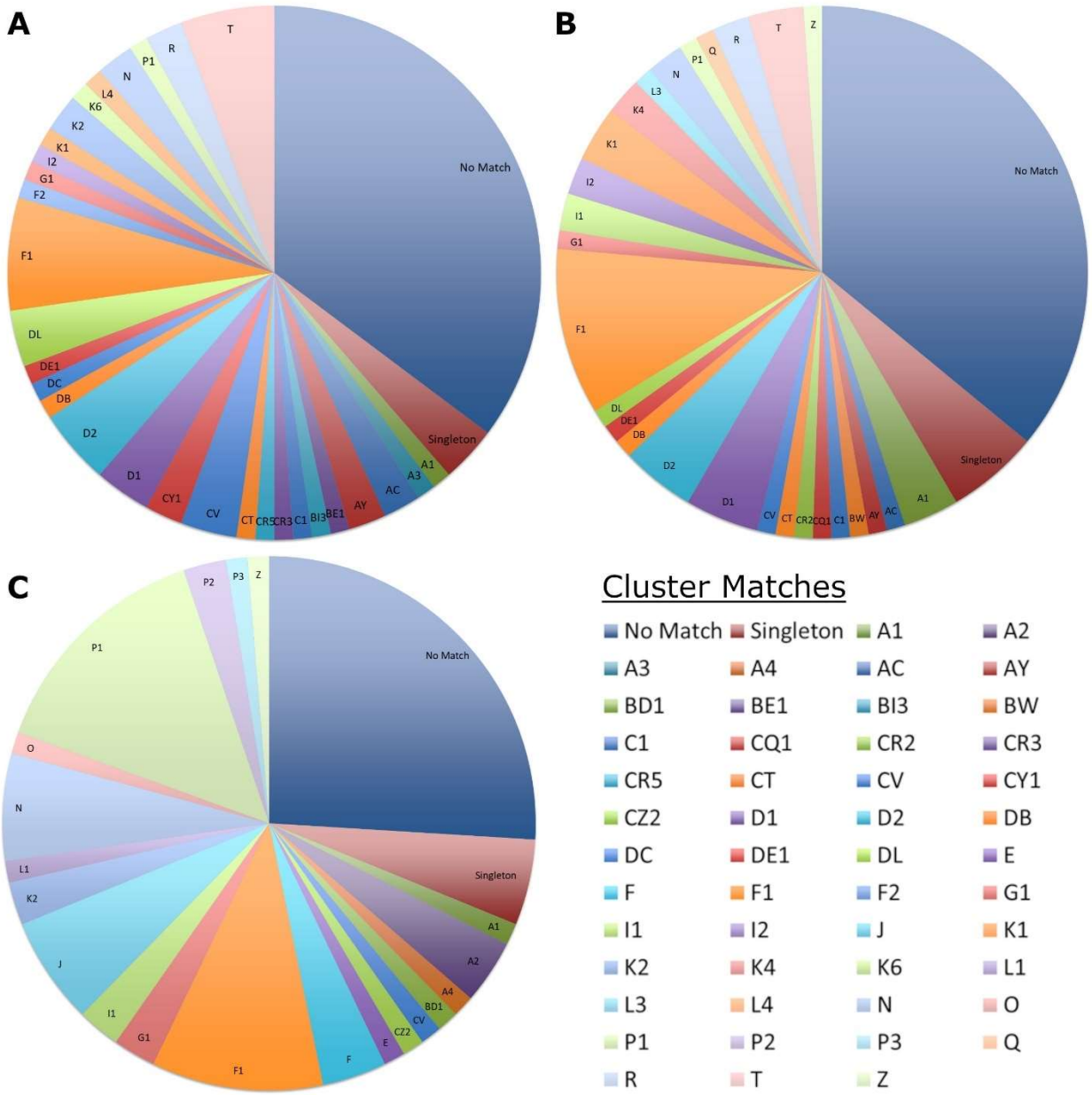

FIG 11 Per-gene cluster matches for a representative putatively intact prophage from each of insertion sites A, B, and C. The matches were identified using a protein BLAST search on the Actinobacteriophage Database (phagesdb.org) (e-value < 1×10^{-4}, bit score > 50). (A) ATCC927_p2, representing insertion site A. (B) ATCC11564_p2, representing insertion site B. (C) ATCC927_p1, representing insertion site C.

No genomes contained multiple prophages with > 50% sequence similarity. The insertion sites tended to be located upstream or downstream of tRNA genes; the insertion site locations in *M. marinum* M and the genes found at each insertion site are listed in Table 11. Attachment (*att*) sites that could be determined by PHASTER are listed in Table 12. Few *att* sites were identified, likely because of the incomplete nature of most of the prophages. Only five MuMC strains were identified by CRISPRFinder as containing putative CRISPR sequences (Table 9).

TABLE 11 Insertion site characteristics

Insertion Site Name	Location in *M. marinum* M genome (bp)	Gene found at insertion site in *M. marinum* M
A	4,812,334	downstream of tRNA-Lys
B	1,856,046	downstream of ssrA tmRNA
C	4,494,467	upstream of tRNA-Asn
D	5,904,289	upstream of tRNA-Phe
E	2,209,878	downstream of competence protein comM
F	2,475,680	upstream of tRNA-Val
G	4,430,818	upstream of tRNA-Met
H	5,470,770	upstream of tRNA-Leu
I	2,430,106	downstream of chitinase/lysozyme precursor *chiA*

TABLE 12 Attachment (*att*) sequences identified by PHASTER

Prophage	Insertion Site	Attachment (*att*) sequence (*attR* and *attL*)
ATCC25039_p1	A	CGACGCGATCGAT
Davis_p1	A	CGACGCGATCGAT
M_p1	A	GATGCGGCCGAGCAGATGCGCCGCAC AGCGTGGCAGCC
M13_p1	A	GATGCGGCCGAGCAGATGCGCCGCAC AGCGTGGCAGCC
ATCC11564_p2	B	CGGGTTCGATTCCCGGCAGCTCCAC
ATCC25039_p2	F	CGGGATAGAGCG
1218R_p2	G	CTCATAATCCGTTGGTCGCGGGTTCGA GCCCCGCCCGCCCCAC
ATCC11564_p3	G	CTCATAATCCGTTGGTCGCGGGTTCGA GCCCCGCCCGCCCCAC

Fig. 12A illustrates the phylogenetic relationships of the 49 MuMC strains. This phylogeny recapitulates the divergence between the two *M. marinum* clades (the M lineage or *M. marinum* subsp. *moffett* and the Aronson lineage or *M. marinum* subsp. *marinum*) and the overall phylogenetic organization of this species described by Das and colleagues based on whole genome comparison (127). This division had initially been described by van der Sar and colleagues based on amplified fragment length polymorphism (AFLP) (128). Strains Davis, ATCC29254 and M13_5434 occupy an intermediate position in the tree, and the striped bass pathogen *M. shottsii* (TM48) occupies a highly divergent branch basal to the Aronson lineage.

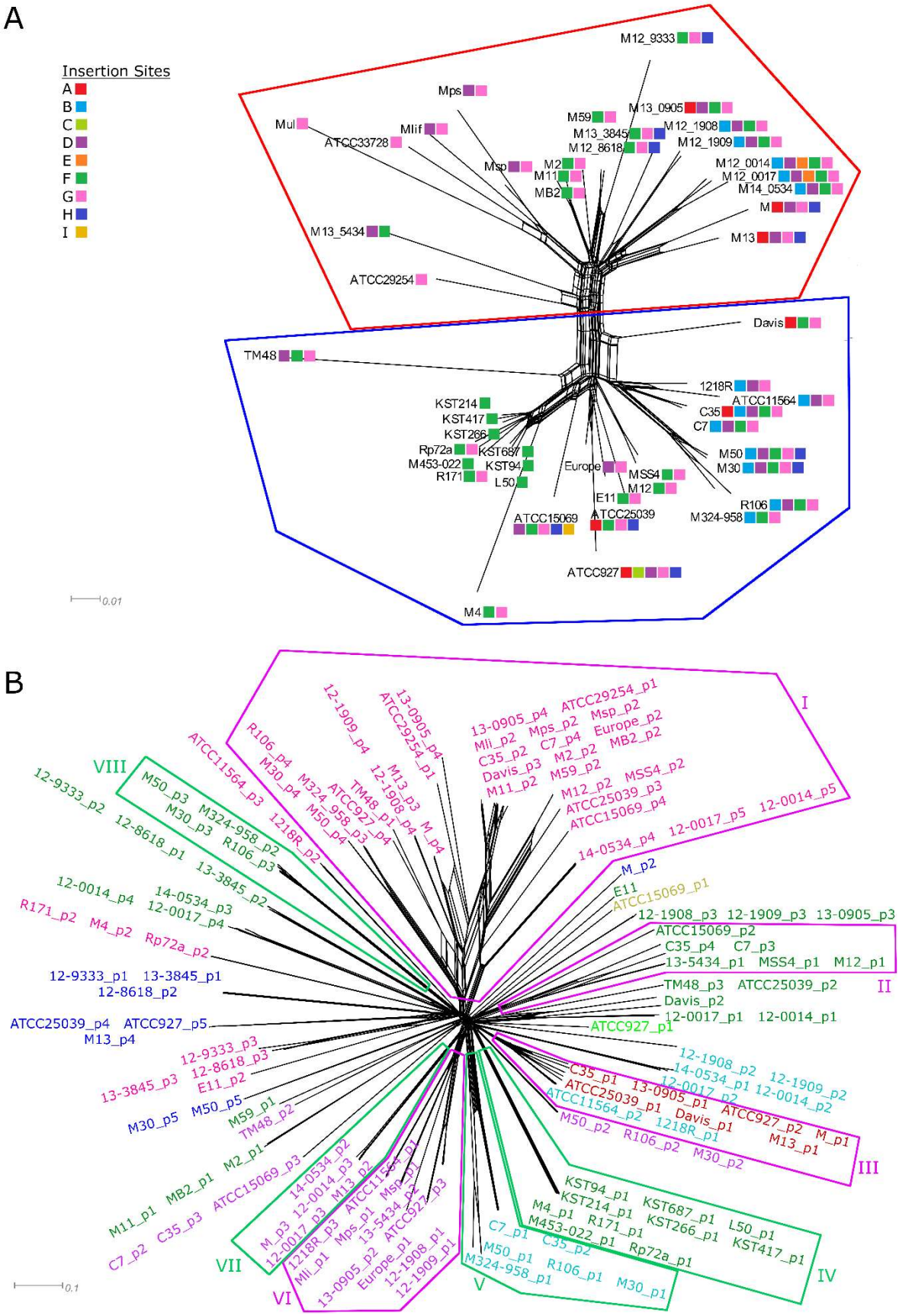
A
Insertion Sites
A
B
C
D
E
F
G
H
I
M12_9333
Mps
M13_0905
M59
Mul
Mlif
M13_3845
M12_1908
ATCC33728
M12_8618
M12_1909
Msp
M2
M12_0014
M11
M12_0017
MB2
M14_0534
M
M13_5434
M13
ATCC29254
Davis
TM48
1218R
KST214
ATCC11564
KST417
C35
KST266
C7
Rp72a
KST687
M50
M453-022
KST94
Europe
M30
R171
L50
MSS4
M12
E11
R106
ATCC15069
ATCC25039
M324-958
ATCC927
M4
0.01
B
I
VIII
II
III
IV
V
VI
VII
0.1

FIG 12 Phylogenetic networks of MuMC strains (A) and of identified prophages (B). The colored boxes in A denote the insertion sites that contain a putative prophage for the given strain, and the small, colored boxes in B denote the insertion site that the given prophage lies within. The red-outlined region in A contains the M lineage strains (*M. marinum* subsp. *moffett*) and the blue-outlined region contains the Aronson lineage strains (*M. marinum* subsp. *marinum*) (128). The purple- and green-outlined regions (labeled I - VIII) in B represent clusters that contain more than three prophages: purple is used for clusters that contain prophages with hosts of both MuMC lineages and green is used for clusters containing prophages from hosts of only one lineage. The figure was generated with the NeighborNet algorithm in the program SplitsTree using the distance matrix from an alignment from the progressiveMauve algorithm in Mauve 2.4.0 (119). The scale bars represent nucleotide substitutions per site.

Fig. 12B demonstrates the predicted phylogeny of the prophages. This splits graph shows that most of the prophages are highly divergent, with branch points near the center of the network. Some of the prophages are arranged in clusters, most of which contain prophages of the same insertion site. Insertion site A prophages are only found in cluster III and insertion site E prophages are only found in a small cluster on the right side (12-0017_p1 and 12-0014_p1). Prophages of insertion sites B, D, F, and H are each found in multiple clusters and singletons. Insertion site B is found in clusters III and V, a small cluster on the right containing 12-1908_p2, and another small cluster on the right containing 12-0017_p2. Insertion site D is found in clusters VI, VII, a small cluster in the lower left containing C7_p2, a small cluster in the lower right containing M30_p2, and a singleton in the lower left (TM48_p2). Insertion site F is very diverse, located in clusters II, IV, and VIII, in addition to four small clusters and five singletons found throughout the splits graph. Insertion site H is found in a singleton in the upper right (M_p2) and three small clusters on the left (containing 12-8618_p2, ATCC927_p5, and M30_p5, respectively). Insertion sites C and I contain only one prophage each, which do not cluster with other prophages.

Some of the clusters contain prophages of closely related MuMC hosts, while others contain prophages of diverse hosts (Fig. 12). The clusters (labeled I - VIII) of more than three prophages in Fig. 12B are outlined in purple or green, based upon if the prophage hosts are from both MuMC lineages (purple) or from only one lineage (green). Many of the clusters contain prophages of closely related hosts of only one MuMC lineage (see clusters IV, V, VII, and VIII), but some have diverse hosts representing multiple MuMC clades in both lineages (see clusters I, II, III, and VI). Cluster III is the

only cluster of more than three prophages that contains prophages of more than one insertion site (A and B). Insertion site A prophages are only found in cluster III while insertion site B prophages are also found in several other clusters. The seven insertion site A prophages in cluster III have diverse hosts from both MuMC lineages, while the two insertion site B prophages in cluster III are from closely related strains. Cluster I contains the largest number of prophages, 32, all in insertion site G, and nine insertion site G prophages are located outside of this cluster.

Phylogenetic trees of individual prophage genes revealed no definite pattern relating to the host phylogeny, as was seen with the network of full prophage sequences (Fig. 13). Also, each gene type had a very distinct topology when compared to the to the trees of the other gene types.

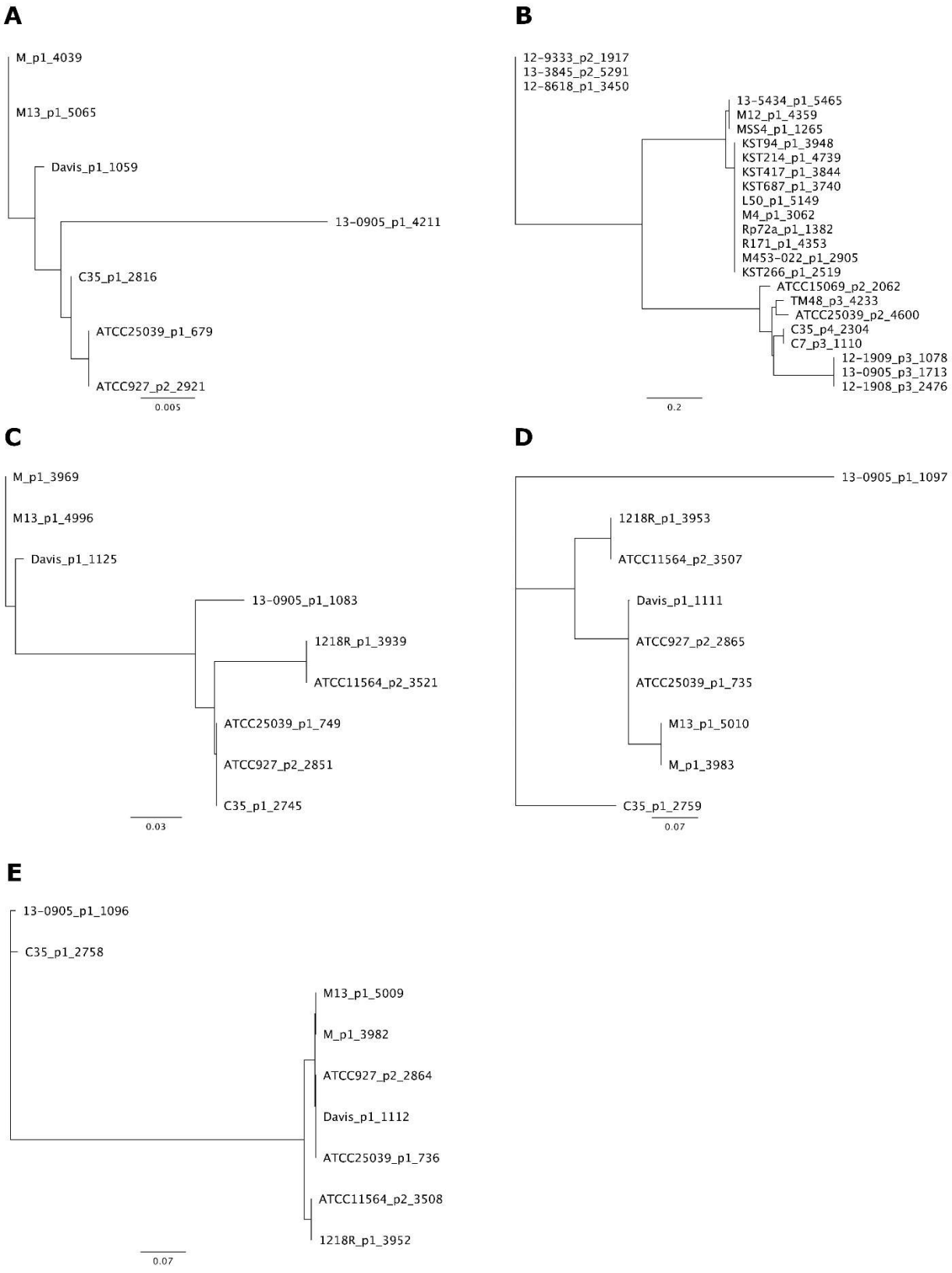
A
M_p1_4039
M13_p1_5065
Davis_p1_1059
13-0905_p1_4211
C35_p1_2816
ATCC25039_p1_679
ATCC927_p2_2921
0.005
B
12-9333_p2_1917
13-3845_p2_5291
12-8618_p1_3450
13-5434_p1_5465
M12_p1_4359
MSS4_p1_1265
KST94_p1_3948
KST214_p1_4739
KST417_p1_3844
KST687_p1_3740
L50_p1_5149
M4_p1_3062
Rp72a_p1_1382
R171_p1_4353
M453-022_p1_2905
KST266_p1_2519
ATCC15069_p2_2062
TM48_p3_4233
ATCC25039_p2_4600
C35_p4_2304
C7_p3_1110
12-1909_p3_1078
13-0905_p3_1713
12-1908_p3_2476
0.2
C
M_p1_3969
M13_p1_4996
Davis_p1_1125
13-0905_p1_1083
1218R_p1_3939
ATCC11564_p2_3521
ATCC25039_p1_749
ATCC927_p2_2851
C35_p1_2745
0.03
D
13-0905_p1_1097
1218R_p1_3953
ATCC11564_p2_3507
Davis_p1_1111
ATCC927_p2_2865
ATCC25039_p1_735
M13_p1_5010
M_p1_3983
C35_p1_2759
0.07
E
13-0905_p1_1096
C35_p1_2758
M13_p1_5009
M_p1_3982
ATCC927_p2_2864
Davis_p1_1112
ATCC25039_p1_736
ATCC11564_p2_3508
1218R_p1_3952
0.07

FIG 13 Phylogenetic trees of individual homologous genes from the prophages analyzed in this study. (A) - (B) integrase genes were grouped into two clusters based on sequence similarity in order to only compare potentially homologous genes. (C) tapemeasure genes. (D) terminase (large subunit) genes. (E) portal genes. The number following the phage name is a unique identifier that represents the specific gene, since in some cases there are multiple genes of a particular type in one prophage. The scale bars represent nucleotide substitutions per site.

A scatterplot of bacterial similarity versus prophage similarity was used to examine relationships between prophage and host similarity. A positive relationship between these measures would be expected if prophage transmission were predominantly vertical. For the possibly-intact prophages in insertion sites A, B, and C, prophage similarity was not observed to have a positive relationship with increasing bacterial similarity (Fig. 14). The Spearman's rank correlation coefficient was -0.15 ($p = 0.50$), and in fact decreased to -0.57 ($p = 0.001$) when considering only the large central region of A/A and A/B comparisons. Therefore, there is no support for a correlation between host similarity and the similarity of the possibly intact prophages, but there is a moderate negative correlation between host similarity and the similarity of the prophages in the central region of the scatterplot, which represents comparisons of insertion site A prophages with insertion site A prophages (A/A) and insertion site A prophages with insertion site B prophages (A/B).

Interesting patterns were observed regarding full prophage similarity and similarities in the integrase genes (Fig. 15). The prophages in insertion sites A and B are somewhat similar, and in fact A/B inter-group identity (59-74% identity) overlaps a great deal with A/A intra-group identity (63-100% identity). Neither are similar to the insertion site C prophage (34-37% identity). Integrase genes did not follow the same pattern (i.e., A and B similar, C as an outgroup); for all three insertion sites, intra-group integrase identity was high (98-100% identity), while inter-group identity was low (25-36%).

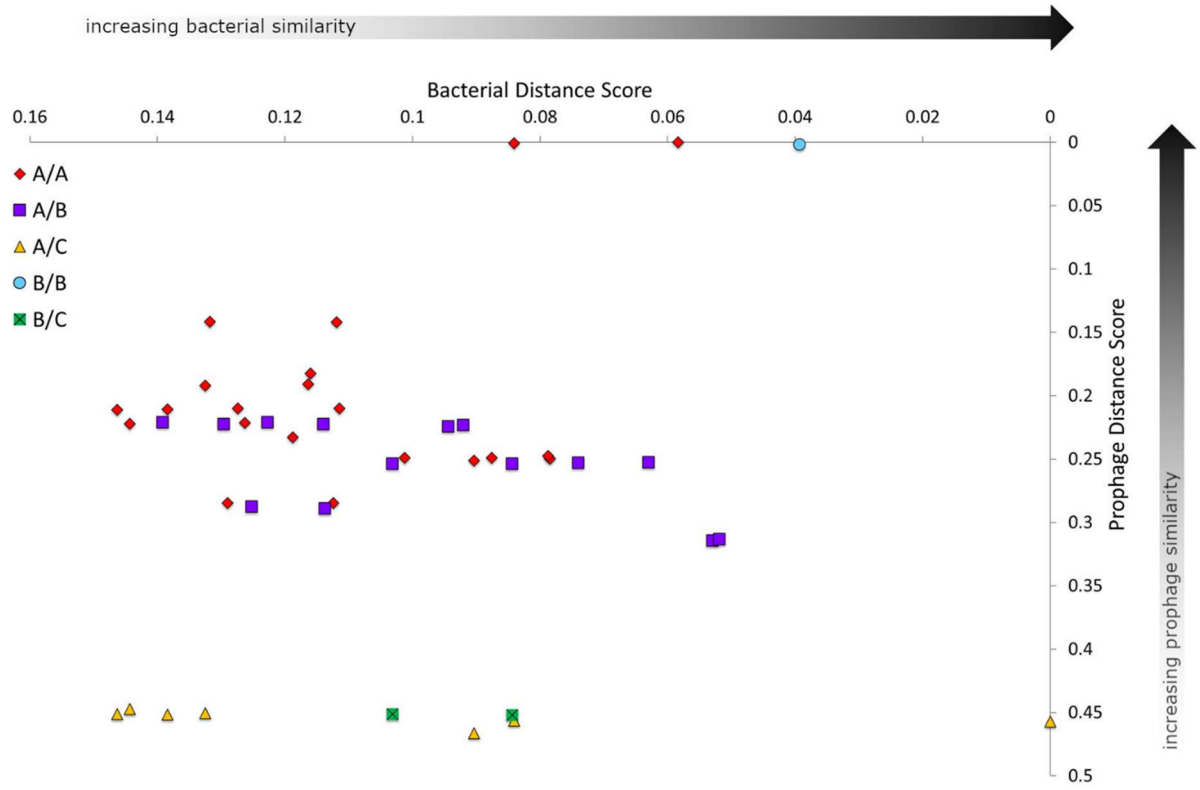

FIG 14 Scatterplot of bacterial distance scores versus prophage distance scores using the possibly intact prophages in insertion sites A, B, and C. Distance scores were generated using the distance matrix from the progressiveMauve algorithm in Mauve 2.4.0 (119). A lower distance score indicates higher similarity (0 represents an exact match, 1 represents no similarities). Each data point represents juxtaposition of genetic distance scores between two prophages (prophage distance score) and distance scores of their hosts (bacterial distance score); the shape and color of the icons represents the insertion site(s) of the compared prophages.

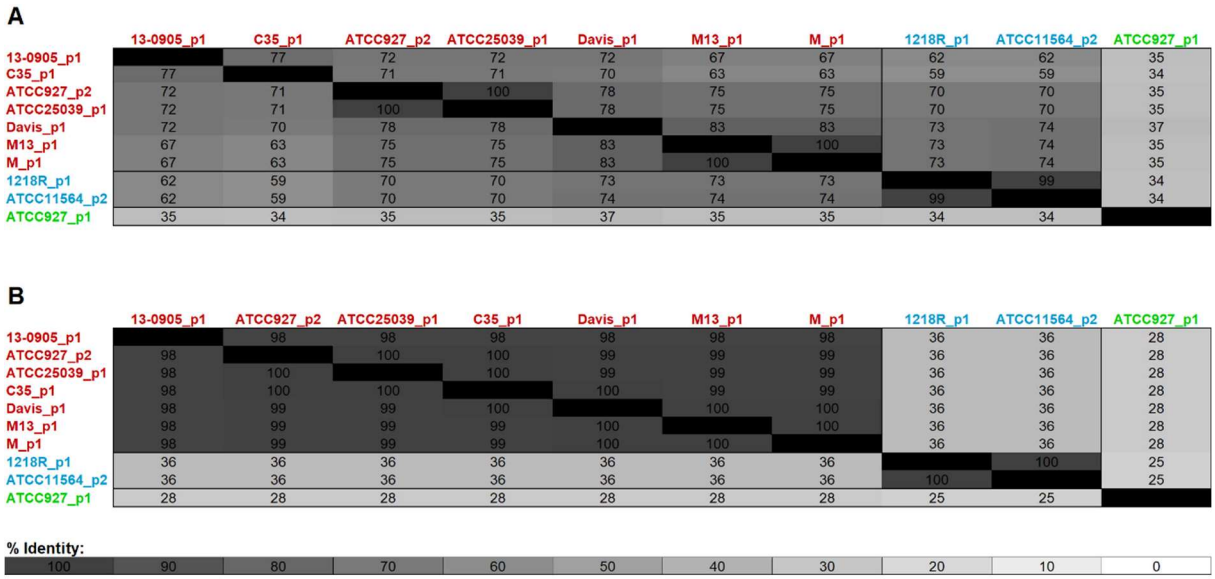

FIG 15 Heatmaps of percent identity of full prophage sequences (A) and integrases (B) for the possibly intact prophages of insertion sites A (red text), B (blue text), and C (green text). Percent identity was calculated with MAFFT in Geneious version 9.1.8 (Biomatters Ltd., Auckland, New Zealand). The darker colors represent higher percent identity.

4.4 DISCUSSION

Prophage phylogenetic analyses can be used to gain a greater understanding of virulence, host-parasite interactions, horizontal gene transfer, evolution, and phage biology. The main goals of the present research were to: 1. Characterize the prophages in 49 MuMC strains and 2. Determine whether the prophages found in these MuMC genomes had been passed vertically from an ancestral strain or had arisen from multiple integration events. The prophages showed evidence of mosaicism, synteny,

superinfection exclusion, and other common features of prophages. Multiple integration events were inferred for most of the prophages, but vertical transmission was possible in the most closely related MuMC strains.

Some of the putative prophages seem to have experienced significant degradation, being composed of as few as one gene (Table 10). These highly degraded prophages were identified due to sharing an insertion site with a larger prophage rather than being identified by PHASTER. There are many more possible prophage genes throughout these genomes that are currently unclear in their origin; integrase genes are especially ambiguous due to their similarity with certain non-phage genes (101). Unfortunately, it may be impossible at this time to be certain of the number of prophage sequences in a genome. The only sequences that can be definitively classified as prophages are those that have been successfully induced, which is not possible for the highly degraded sequences that comprise most prophages (16).

Many of the findings are consistent with well-known concepts of phage biology. Although many of the genomes contained multiple prophages, no instances were identified of multiple prophages sharing more than 50% sequence similarity within a single bacterial strain, as expected due to superinfection exclusion (100). The overwhelming majority of identified prophages were incomplete, consistent with the view that prophage degradation generally occurs rapidly after integration, unless a gene valuable to the host is provided by the prophage. Most of the insertion sites are located at tRNA sequences, which is a common arrangement, especially for integrases of the tyrosine recombinase class (129) (Table 11). PHASTER was able to identify *att* sites in some of the prophages, which are the remnants of the recombination sites of the phage

(*att*P) and the bacterium (*att*B), providing support for those sequences being authentic prophages (Table 12). The prophages were found to be highly mosaic in their genetic structure; none of the prophages could be assigned to pre-existing phage clusters on the Actinobacteriophage Database, even though many of the prophage genes were similar to genes in that database (Fig. 11). This mosaicism suggests that frequent recombination events occurred in these phages. However, synteny was observed in the order of gene types for the likely intact prophages of insertion sites A, B, and C (Fig. 10).

CRISPR (Clustered regularly interspaced short palindromic repeats) genes, along with *cas* (CRISPR-associated) genes, make up an adaptive immune mechanism in bacteria; DNA from invading phages is cleaved and incorporated into the CRISPR sequence itself, allowing the CRISPR-Cas system to recognize and destroy similar invaders in the future (130). Although only five of the MuMC genomes contained likely CRISPR sequences, precluding the drawing of any definite conclusions from this data, it was interesting that the strains with CRISPR sequences often contained a higher than average number of prophages (Table 9). This contrasts with other studies that have found fewer prophages in CRISPR-containing strains (130-132); however, some studies have been inconclusive (133). Multiple factors likely influence the number of prophages found in a bacterial strain, complicating the search for a relationship between CRISPR and prophage prevalence.

MuMC are divided into two general clades, which were initially designated the M lineage (including a branch encompassing the *M. ulcerans* ecovars) and the Aronson lineage by van der Sar and colleagues based on AFLP analysis (128). Das and

colleagues used whole genome analysis to support the distinctness of these two branches and suggested that the M lineage be named *M. marinum* subsp. *moffett* and the Aronson lineage be named *M. marinum* subsp. *marinum* (127). They also found that much of the high genomic diversity of *M. marinum* can be attributed to prophage regions, although the prophages were variable within each lineage and could not predict the lineage of a particular strain (127). In many bacterial species, strain-specific differences have been demonstrated to be the result of prophages (100). The present research supports these findings; prophage composition varied greatly among the analyzed strains but could not predict host relatedness except for the most recently diverged strains.

Splits graphs, such as those produced by the program SplitsTree, are a valuable way to illustrate phylogenies in which multiple branching patterns are supported by the input data; this method is especially well-suited for representing the complexity of microbial evolutionary histories (121). Each node of the splits graph shows different possible splits simultaneously, forming a box-shaped pattern at the node, with the length of each edge of the box representing the weight of the support for that pattern (121). If all of the splits of a splits graph were fully compatible, then there would only be one split per node and the figure would look identical to a traditional phylogenetic tree; if the splits were instead all weakly compatible, a more net-like arrangement would be seen, as there would be a greater abundance and length of edges (121). The topology of the splits graph generated for the prophages did not recapitulate that for the bacterial hosts (Fig. 12). Multiple integration events were inferred for most of the prophages analyzed, based on the findings that none of the prophages were found among all

members of deep MuMC clades. The only cases in which a single integration event was inferred for prophages in multiple strains were in cases of closely related strains with a recent common ancestor. For example, for the groupings of closely related strains ATCC11564 and 1218R; M11, M2, and MB2; and 12-1908 and 12-1909, each grouping contains similar prophages in the same insertion sites. There are some cases where differences are seen in closely related strains, however. Strains C7 and C35 share similar prophages in insertion sites B, D, F, and G, but C35 also contains a prophage in insertion site A (and is the only case in which one strain contained both an insertion site A prophage and an insertion site B prophage). The prophages of strains M and M13 cluster for insertion sites A, D, and G, but their small prophages in insertion site H are not similar. Strains M4, M453-022, R171, Rp72a, L50, and all the KST strains all contain prophages in insertion site F which form a cluster in the prophage splits graph; however, M4, R171, and Rp72a also contain a small prophage in insertion site G. M30, M50, M324-958, and R106 all have similar prophages for each of insertion sites B, F, and G, but differ in whether they contain prophages at insertion sites D and H. The insertion site D prophages in M30, M50, and R106 each are composed of 44 CDS, while the H prophages in M30 and M50 are each composed of only eight CDS. When a strain contains a large prophage that a closely related strain does not have, the prophage may have integrated recently relative to the diversification of those strains. However, when the difference between closely related strains is a small prophage sequence, there are a few possibilities. The prophages may be degrading rapidly, the designation of these sequences as prophages may be erroneous, or these may simply

be very small phages or satellite phages (small phages that parasitize larger phages) (134).

Because the only evidence for ancient shared integration events within these strains occurred in very shallow nodes, most of the integration events are inferred to be relatively recent compared to the diversification of these MuMC strains. Phylogenetic trees of individual prophage genes recapitulated neither the phylogenetic network of full prophage sequences nor the other single-gene trees, demonstrating the value of full sequence comparison when analyzing phages and prophages (Fig. 13).

In addition, if the prophages had predominantly arisen from integration events ancestral to the diversification of these *M. marinum* strains, one would expect to see prophage similarity increase with host strain similarity. No relationship between host similarity and prophage similarity was found (Fig. 14). Although using possibly intact prophages for this analysis decreases the likelihood of older integration events being included, it also provides the highest likelihood that only authentic prophages are included in the analysis. The Spearman's rank correlation coefficient for bacterial similarity vs. possibly intact prophage similarity was close to zero and not statistically significant, and in fact for the large central region (A/A and A/B comparisons) it had a statistically significant moderate negative correlation. The central region excluded the low-prophage similarity comparisons between the insertion site C prophage and the insertion site A and B prophages (which are not likely to be homologous due to differences in both insertion site and full prophage similarity), and the high-prophage similarity comparisons between ATCC927_p2 and ATCC25039_p1, M13_p1 and M_p1,

and ATCC11564_p2 and 1218R_p1 (identical or almost-identical prophages in closely related hosts).

Prophages did not appear to be vertically inherited, even in the case of similar prophages in the same insertion site, except in the most closely related strains. This conclusion is consistent with results of other mycobacterial prophage phylogeny studies. Fan and colleagues concluded that prophage similarity was not a good predictor of overall genome similarity after comparing a prophage phylogenetic tree to a tree of the host genomes in 23 species of mycobacteria (108). Sassi and colleagues analyzed 48 genomes of *M. abscessus* and found that, based on prophage sequence similarity, the prophages tended to cluster according to host subspecies, although multiple clusters and singletons were present for each subspecies (110). These studies reached different conclusions; however, the sets of genomes used in the studies differed considerably. In the former study, the genomes used represented the diversity of the genus, including such species as *M. canettii*, *M. avium*, *M. abscessus*, and *M. smegmatis*, and only one to five strains of each species (108). The latter study used many strains of one species, which the phylogenetic tree suggests have undergone recent diversification (110). Although the overall topology of the *M. abscessus* prophage phylogenetic tree differed from that of the host tree, the recently diverged strains do cluster together in the prophage tree, indicating that prophages may have been vertically inherited in these closely related strains (110). These studies and the present research support the idea that independent integration events may be the most likely scenario for prophages in different species and in more distantly related strains, but vertical inheritance appears to be more common for similar prophages of strains with a more recent divergence.

Interesting patterns in integrase identity were revealed using heatmaps and percent identity scores (Fig. 15). The prophages in insertion sites A and B are almost as similar (average percent identity 68.8, standard deviation ± 5.7) as insertion site A prophages are to each other (average percent identity 75.6, standard deviation ± 9.6), while the prophage in insertion site C is not similar to either of the other evaluated groups (average percent identity 34.9, standard deviation ± 0.9). By comparing the integrase genes of prophages sharing an insertion site versus prophages in different insertion sites, it can be determined whether there is a relationship between insertion site and integrase sequence. Despite the possibly intact prophages in insertion sites A and B being similar, their integrase genes are dissimilar (average percent identity 36.0, standard deviation 0). Similarly, Vale and colleagues found that in *Helicobacter pylori* prophages, the population of the prophage (clustering based on integrase and portal genes) was not a good predictor of overall prophage sequence similarity (106). Bobay and colleagues analyzed *Escherichia* and *Salmonella* prophages and determined that integrases with high similarity tend to be specific for the same integration site, while integrase similarity does not correlate as strongly with prophage type or even host genus (102). Rezaei Javan and colleagues analyzed both prophages and satellite prophages and found that integrases with high sequence similarity tend to be found in the same insertion sites (134). These observations support the notion that integrases are insertion site-specific and can be found among diverse phages. While integrase similarity may indicate that two phages or prophages insert at a similar nucleotide sequence, it should not be used as an estimate of overall phage similarity.

In-depth analyses of prophage phylogeny are invaluable, especially those utilizing large genome libraries. The present research contributes to the growing knowledge of mycobacterial prophage genetics, phylogeny, and evolution. In future research it would be beneficial to investigate the phylogeny of prophages in other species of aquatic mycobacteria, and to compare these results to the MuMC data. It is hoped that others will perform similar large-scale analyses of prophage phylogeny on other bacterial genera so that comparisons can be made. As phages become increasingly utilized in medicine, industry, research, and education, it becomes more essential than ever to understand the interaction between these parasites and their hosts.

CHAPTER 5

MYCOBACTERIOPHAGE HOST RANGE STUDIES IN MULTI-WELL PLATES

5.1 INTRODUCTION

Bacteriophages (phages), viruses that infect bacteria, can be used to treat bacterial infections, a practice known as phage therapy (31). One of the benefits of phage therapy is host specificity; phages can only infect hosts within the phage's host range (33). Host specificity is, in part, determined by the receptor binding protein(s) (RBP) of the phage, which bind to bacterial membrane surface receptors during the adsorption stage of infection (11). The surface receptors can be proteins, carbohydrates, or lipids; common receptors include peptidoglycan, teichoic acids, polysaccharides, and lipopolysaccharide (LPS) (15). RBPs are found on the phage tail fibers, tail spikes, and/or baseplate (15). For temperate phages which can integrate their genetic sequence into the host genome, the integrase gene can affect host specificity because integrases are often specific to a locus which must be present in the host genome (16).

Phages destined for phage therapy can originate from phage banks, biotechnology companies, or academic laboratories (32). However, sometimes it is difficult to find phages that infect a particular host. No obligately lytic phages have been cultured that infect *Clostridium difficile* (35) or *Bifidobacterium spp.* (36). Some phages require glycine (40), a low agar concentration (34), ampicillin (41), or mucins (42) to produce clearly visible plaques. Metagenomics and microscopy techniques indicate that

cultured phages make up an unknown, but likely small, fraction of phages present in the environment (22). Current isolation techniques may not be appropriate for every phage.

The genus *Mycobacterium* comprises almost 200 species, including obligate pathogens, opportunistic pathogens, and saprophytes (44). Tuberculosis, caused by *M. tuberculosis*, is one of the deadliest human diseases; in 2016, 1.7 million people died from tuberculosis-related causes and it has been estimated that 23% of humans are latent *M. tuberculosis* carriers (49). *M. tuberculosis* is difficult to treat and prone to developing antibiotic resistance (49). Nontuberculous mycobacteria (NTM) are found in soil, water, and animals (51). Many NTM are opportunistic pathogens, especially in immunocompromised individuals (52). Phage therapy has been used to treat *M. abscessus* subspecies *massiliense* with considerable clinical improvement as a result (48), and phages have been investigated for the treatment of *M. tuberculosis*, *M. bovis, M. avium, and M. ulcerans* infections (47). Cultured phages also exist for *M. fortuitum* (53), *M. phlei* (53), *M. vaccae* (53), *M. chelonae* (55), and *M. scrofulaceum* (54). Rybniker and colleagues investigated the host range of 14 phages in 17 mycobacterial strains of the species *M. tuberculosis*, *M. bovis*, *M. avium*, *M. scrofulaceum*, *M. ulcerans*, *M. marinum*, *M. fortuitum*, *M. chelonae*, and *M. smegmatis*; every strain could be infected by at least three of the phages, except for the MuMC strains which were not found to be infected by any of the phages (54). Prophages have been induced from *M. marinum* (2), *M. avium* (135), *M. intracellulare* (135), *M. scrofulaceum* (135), and *M. fortuitum* (56). Compared to the small number of isolated phages for most mycobacterial species, there are hundreds of sequenced phages available for *M. smegmatis* as a result of the Science Education Alliance–Phage Hunters Advancing Genomics and

Evolutionary Science (SEA-PHAGES) program (27), which can be accessed via the Actinobacteriophage Database (phagesdb.org) (92).

The aims of the present study were to compare the traditional plaque assay technique with the multi-well plate turbidity assay for host range analysis, and to evaluate whether any pre-purified phages can infect *M. ulcerans*-*M. marinum* complex (MuMC), *M. fortuitum*, and/or *M. chelonae*. These species can cause systemic granulomatous inflammation in fishes and zoonotic infections in humans (1). Genetic analyses were used to select phages from a pre-existing phage collection for host range studies in the three species. Ten phages were selected, then screened by plaque assay and multi-well plate turbidity assay, using undiluted phage stock solutions. Phage-host combinations were selected and used for plaque assays and multi-well plate turbidity assays with dilution series of the phages.

5.2 MATERIALS AND METHODS

Bacterial Strains

Table 13 shows the bacterial strains used in this study. *M. smegmatis* mc^2155 was only used in the screening and in plaque assays to ascertain phage titers, since it was already known that these phages could infect that bacterium.

TABLE 13 Bacterial strains

Species	Strain
M. smegmatis	mc^2155
M. chelonae	324-818
	M3
M. fortuitum	M5
	M6
M. marinum	ATCC927
	ATCC11564
	KST214
	M
	M2

Each strain was grown by combining 13.35 mL Middlebrook 7H9 broth (7H9), 10 mL albumin dextrose (AD) supplement (for *M. smegmatis*) or oleic albumin dextrose catalase (OADC) supplement (for *M. chelonae*, *M. fortuitum*, and *M. marinum*), 1 mL 100 mM $CaCl_2$, 100 µL 50 mg/mL carbenicillin (CB), 100 µL 10 mg/mL cyclohexamide (CHX), and 0.05% Tween 80 (polyoxyethylene sorbitan monooleate). The cultures were incubated at 37°C (for *M. smegmatis*) or 30°C (for *M. chelonae*, *M. fortuitum*, and *M. marinum*), shaking at 140 rpm, until reaching approximately 10^6 CFU/mL (estimated by turbidity, then confirmed by colony count). Cultures were centrifuged at 2,000 x *g* for 10 minutes to pellet the bacteria, the supernatant was removed, and the pellet was resuspended in fresh medium (without Tween) before use.

Purified Phages

The pre-purified phages used in this study were obtained from the laboratory of Dr. Graham Hatfull at the University of Pittsburgh and had been selected from the Actinobacteriophage Database (phagesdb.org). All phages were in the Order *Caudovirales*, Family *Siphoviridae*. Characteristics of these phages can be seen in Table 14.

TABLE 14 Characteristics of the phages used in this study. Data concerning lifestyle, cluster, subcluster, and isolation host were retrieved from the Actinobacteriophage Database at phagesdb.org (lifestyle, cluster, and subcluster designations are based on genetic data)

Name	Cluster	Subcluster	Lifestyle	Isolation Host
MalagasyRose	Singleton	[none]	Unknown	*M. smegmatis* mc²155
Power	A	A2	Temperate	*M. smegmatis* mc²155
Cucurbita	CQ	CQ1	Temperate	*Gordonia terrae* 3612
Schmidt	CU	CU4	Temperate	*Gordonia terrae* 3612
Hawkeye	D	D2	Lytic	*M. smegmatis* mc²155
Noely	FE	[none]	Lytic	*Arthrobacter globiformis* B-2979
Shaobing	K	K1	Temperate	*M. smegmatis* mc²155
Omnicron	K	K5	Temperate	*M. smegmatis* mc²155
Rando14	K	K5	Temperate	*M. smegmatis* mc²155
Mendokysei	T	[none]	Temperate	*M. smegmatis* mc²155

Six phages were selected based on their results from two methods of genetic analysis. Four phages were selected without the genetic analyses to evaluate whether the genetic analyses improved the chances of finding a phage that could infect the bacteria of interest.

Prophage genes of MuMC were analyzed for similarity to genes of cultured phages. The Actinobacteriophage Database protein BLAST (phagesdb.org) was used to find the closest match to each gene in the prophages (e-value < 1×10^{-4},bit score > 50). The purified phages with matches to the prophage genes were further analyzed with HostPhinder, which compares the genetic similarity (using k-mers) between the input phage and phages known to infect various hosts (136). Several mycobacterial hosts are found in the HostPhinder database used for determining possible phage hosts (136). The phages Schmidt and Hawkeye were selected based on BLAST results; Power, Cucurbita, and Shaobing were selected based on HostPhinder results; and Mendokysei was selected based on both methods (Table 15). The other four phages were selected to represent a range of genetic diversity. MalagasyRose is a singleton (not similar enough to any other phages to be assigned to a cluster). Noely was isolated with *Arthrobacter globiformis*, an isolation host not represented by the other selected phages. Omnicron and Rando14 are part of cluster K, which is known to contain phages with a broad host range (Power is in subcluster A2, which is also known to have a broad host range) (137). Phages of the same cluster or subcluster often have similar host ranges, as demonstrated by Jacobs-Sera and colleagues (137) and it was hypothesized that the cluster K phages used in this study would share a host range, while the other phages would have unique host ranges.

TABLE 15 Results of genetic analyses. The protein BLAST search was performed on phagesdb.org with the protein sequences from MuMC prophages against a database of actinobacteriophage proteins (e-value < 1×10^{-4}, bit score > 50)

Phage Name	Protein BLAST	HostPhinder
MalagasyRose	[none]	*M. smegmatis*
Power	Integrase	*M. smegmatis*, *M. tuberculosis*
Cucurbita	Hypothetical protein	*M. smegmatis*, *M. tuberculosis*, *Caulobacter crescentus*, *Gordonia rubripertincta*, *Rhodococcus equi*
Schmidt	Integrase	*M. smegmatis*, *Caulobacter crescentus*
Hawkeye	Minor tail protein, tapemeasure, hypothetical protein	*M. smegmatis*
Noely	[none]	*M. smegmatis*, *Tsukamurella paurometabola*
Shaobing	Integrase, hypothetical protein	*M. smegmatis*, *M. tuberculosis*, *M. avium*, *Streptomyces venezuelae*
Omnicron	[none]	*M. smegmatis*
Rando14	DNA primase/polymerase	*M. smegmatis*
Mendokysei	Lysin A, terminase small subunit, excisionase, exonuclease, hypothetical protein	*M. smegmatis*, *M. tuberculosis*, *Burkholderia cenocepacia*, *Tsukamurella paurometabola*, *Caulobacter crescentus*, *Rhodococcus equi*

Statistical Analysis

For 96-well plate turbidity assays with dilution series, SPSS (Statistical Package for the Social Sciences) was used to perform a one-way Analysis of Variance (ANOVA) on each time point for each dilution of each phage in the 96-well plate assays. A one-tailed Dunnett's test was used for post-hoc testing (65). For 96-well plate turbidity screenings (undiluted phage) a t-test was used. A p-value of less than 0.05 was considered statistically significant.

Host Range Screening

The selected phages were subjected to a screening step to conserve resources. The following criteria identified the phage/host combinations selected for further testing. Phages with statistically significant results (see Statistical Analysis above) on the screening (non-dilution) turbidity assay were considered for further analysis. If more than two phages fulfilled this first criterion, the two phages with the highest number of days of statistically significant results were selected for each strain. If more than two phages produced statistically significant results on the same number of days, the ones with the lowest p-values were selected. *M. smegmatis* pairings were not selected for further testing, as these phages were already known to infect that bacterial strain.

The turbidity assay screening was performed in a manner similar to the typical multi-well plate turbidity assay (see Multi-Well Plate Turbidity Assay below), except with only three wells per phage (undiluted). The plaque assay screening was performed similarly to the typical plaque assay (see Plaque Assay below), but with undiluted phage. Because high titers of phage can sometimes kill bacteria without a productive

lytic cycle (lysis from without) (66), the phage/host combinations with the most promising results from the screening were used in further testing with dilution series.

Plaque Assay

Plaque assays were performed according to the Phage Discovery Guide (25) on Luria agar Petri plates with a 0.4% agar overlay. Phage solutions were spotted onto the plates (5 µL per spot) in triplicate. The plates were incubated at 37°C (*M. smegmatis*) or 30°C (*M. chelonae*, *M. fortuitum*, and *M. marinum*) until the bacterial lawn had formed and observations were made each day.

Multi-Well Plate Turbidity Assay

Dilutions of pre-purified phage (20 µL per well of dilutions starting at 10^8 PFU/mL) mixed with host bacteria (10 µL per well of approximately 10^6 CFU/mL) were used in the turbidity assays. This is equivalent to approximately 1 x 10^4 bacterial cells per well mixed with a range of zero to 2 x 10^6 phages. Titers were determined based on spot tests on *M. smegmatis* mc^2155.

A plaque assay was performed on phage solutions on *M. smegmatis* mc^2155 prior to phage detection experiments so the appropriate concentrations could be estimated. Eight dilutions of each phage solution (beginning with 10^8 PFU/mL) were used along with a control (phage buffer), each performed in triplicate. To each well of the 96-well plate was added 148 µL 7H9 broth, 20 µL AD supplement (for *M. smegmatis*) or OADC supplement (for *M. chelonae*, *M. fortuitum*, and *M. marinum*), 2 µL 100 mM $CaCl_2$, 0.2 µL CB, 0.2 µL CHX, 10 µL bacterial culture, and 20 µL phage

solution (or phage buffer control solution). The absorbance at 595 nm was measured in a plate reader (BMG Labtech, Cary, NC) directly after mixing the solution and every 24 hours afterwards for seven days. Because it was found that some of the phages exhibit a rapid decrease in titer during storage, another plaque assay was performed at the time of starting the phage detection experiment and read the next day. Corrections were made to the starting concentrations to account for phage titers being less than expected based on the initial plaque assays. A statistically significant difference ($p<0.05$) in absorbance between the control group and at least three dilution groups on at least two consecutive days was considered a positive result.

5.3 RESULTS

The screening resulted in positive plaque assays and turbidity assays for all phages on *M. smegmatis* mc^2155, as expected. The plaque assay screening was also positive for several phages on the *M. fortuitum* strains (Table 16). No individual plaques were seen for the *M. fortuitum* strains and the areas of clearing were turbid and faint. Many of the phages that produced areas of clearing on the plaque assay screening on an *M. fortuitum* strain did not produce positive results on the turbidity assay screening for the same strain (Power, Cucurbita, Schmidt, Hawkeye, Shaobing, Omnicron, Rando14, and Mendokysei). *M. fortuitum* M6 with Power was the only pairing which had positive results on both the screening plaque assay and the screening turbidity assay, outside of the *M. smegmatis* pairings. None of the MuMC strains had positive results on the plaque assay screening, although many had positive results on the turbidity assay screening (Power, Cucurbita, Schmidt, Noely, Shaobing, and Rando14).

TABLE 16 Screening results. Cells are blank when no statistically significant result exists. P=positive for plaques on plaque assay, p = clearing but no individual plaques on plaque assay, T=positive for turbidity assay in 96-well plates. Shaded cells with bold text indicate that the phage/host combination was selected for further analysis

Bacterial Strains	Malagasy Rose	Power	Cucurbita	Schmidt	Hawkeye	Noely	Shaobing	Omnicron	Rando14	Mendokysei
	Phages									
M. smegmatis mc^{2}155	P, T	P, T	P, T	P, T	P, T	P, T	P, T	P, T	P, T	P, T
M. chelonae M324-818										
M. chelonae M3					**T**					
M. fortuitum M5		p	p	p			p	p	p	p
M. fortuitum M6		**p, T**	p	p			p	**T**		
M. marinum ATCC927			**T**			**T**			T	
M. marinum ATCC11564			**T**				**T**			
M. marinum KST214										
M. marinum M			**T**				**T**			
M. marinum M2		**T**		T		T	T		**T**	

Eleven phage-bacterium pairings were selected from the screening for further analysis, which included the final plaque assay and final turbidity assay using serial dilutions of the selected phages. None of the selected pairings produced positive dilution plaque assays, except for *M. fortuitum* M6 with Power. This pairing produced a turbid area of clearing only at the highest phage concentration (10^8 PFU/mL); no individual plaques were observed and the dilution turbidity assay was not positive. Only three phage-host pairings had positive outcomes for the dilution turbidity assay (a

statistically significant difference in absorbance between the control group and at least three dilution groups on at least two consecutive days): Hawkeye with *M. chelonae* M3, and Noely and Cucurbita with *M. marinum* ATCC927 (Fig. 16).

5.4 DISCUSSION

Despite the importance of phages in research, education, biotechnology, and medicine (6), and the demand for therapeutic phages capable of infecting each pathogen of interest (31), only a small percentage of phages have been cultured (22). In the present research, the traditional plaque assay technique was compared to the multi-well plate turbidity assay in host range analyses. If phages have different culture requirements or are more likely to complete a productive infection cycle under certain conditions, it is important to develop alternate techniques for host range analysis. It is possible that phages of aquatic bacterial hosts are adapted to adsorbing onto their hosts in liquid medium, in which case multi-well plate techniques may be preferable to the traditional plaque assay method of host range determination for these phages.

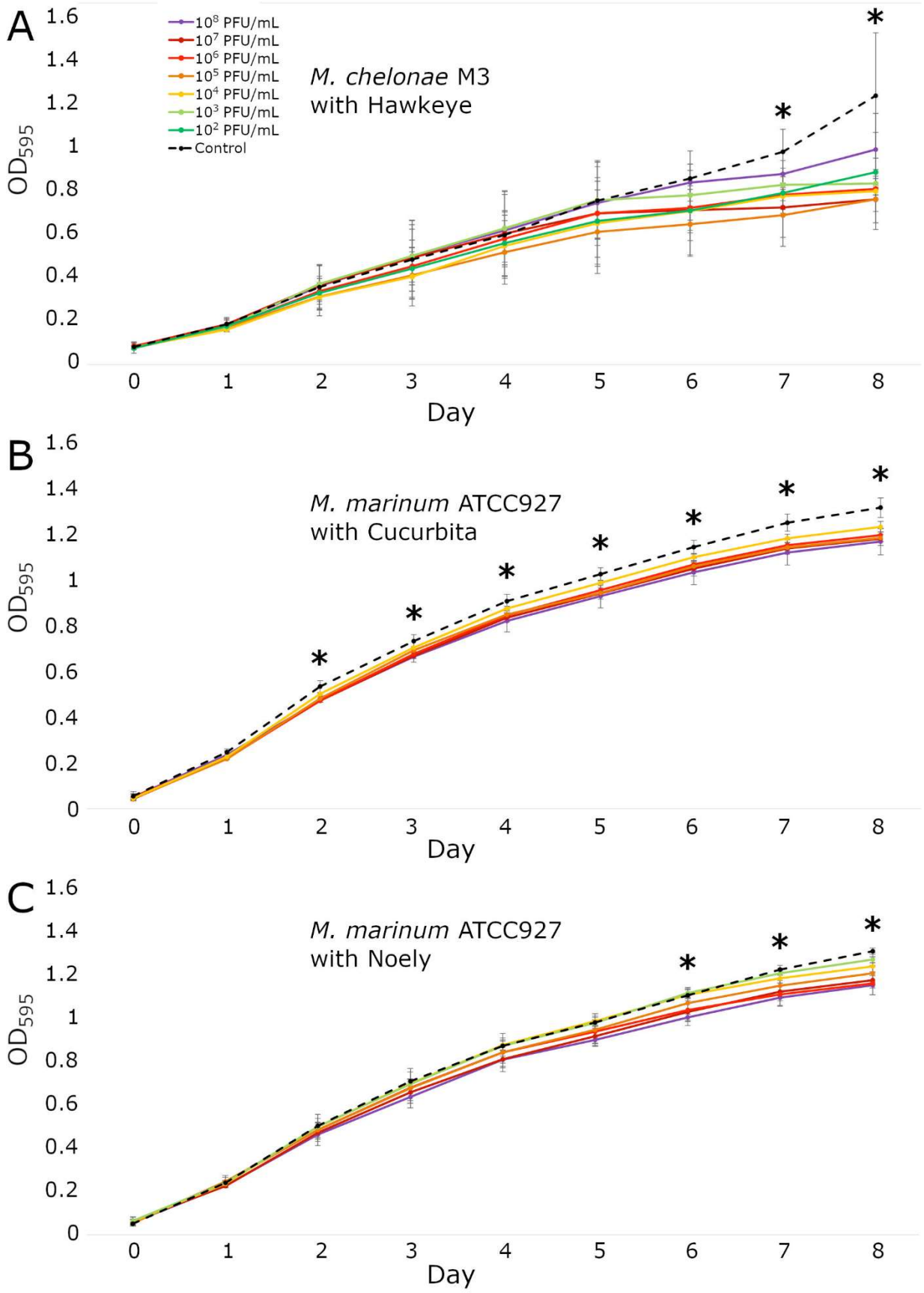
A
B
C
10^8 PFU/mL
10^7 PFU/mL
10^6 PFU/mL
10^5 PFU/mL
10^4 PFU/mL
10^3 PFU/mL
10^2 PFU/mL
Control
M. chelonae M3
with Hawkeye
M. marinum ATCC927
with Cucurbita
M. marinum ATCC927
with Noely
OD_{595}
Day
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
0
1
2
3
4
5
6
7
8
*

FIG 16 Turbidity assays with positive outcomes. (A) *M. chelonae* M3 with Hawkeye, (B) *M. marinum* ATCC927 with Cucurbita, (C) *M. marinum* ATCC927 with Noely. In (A), statistically significant results were 10^7 PFU/mL-10^4 PFU/mL on day 7 and 10^7 PFU/mL-10^2 PFU/mL on day 8. Asterisks indicate days on which there was at least on dilution with a statistically significant result. In (B), statistically significant results were 10^8-10^5 PFU/mL on days 2-6, and 10^8-10^4 PFU/mL on days 7-8. In (C), statistically significant results were 10^8 PFU/mL on day 6, 10^8-10^6 PFU/mL on day 7, and 10^8-10^5 PFU/mL on day 8. Error bars represent +/- 1 standard deviation of triplicate wells.

Outside of the *M. smegmatis* pairings, the plaque assay and turbidity assay results were difficult to interpret and did not necessarily indicate productive phage infections. However, since a statistically significant decrease in growth rate was observed for some of the turbidity assays and decreased bacterial lawn formation were seen in some of the plaque assays, it is possible that some of these phages were able to initiate infection with some of the bacterial strains, even if the infection was ultimately incomplete. It is less likely that the phages were inhibiting bacterial growth without any host specificity, since the set of phages producing positive results differed based on the bacterial strain, and because more genetically similar strains often produced similar screening results. For example, *M. marinum* strains M and M2 are more genetically similar to each other than to the rest of the *M. marinum* strains, and they produced identical results on the phage screening (Table 16).

Eight of the ten tested phages produced areas of clearing on one or both *M. fortuitum* strains in the plaque assay screening, and none of the phages produced any clearing or plaques on the *M. chelonae* or MuMC strains in the plaque assay screening (Table 16). One of these phage-host pairings (*M. fortuitum* M6 with Power), which was also positive for the turbidity assay screening, was selected for the dilution plaque assay and dilution turbidity assay. Positive results were not seen in the dilution turbidity assay, and the dilution plaque assay resulted in clearing only for the highest concentration of the phage. Individual plaques were not seen. Bacterial lysis only at high phage titers is often attributed to lysis from without, which will not result in individual plaques at any phage dilution (66). This phenomenon can occur when adsorption of a high density of phages occurs on the bacterium; a productive phage infection is not completed (66).

Other causes of bacterial clearing without individual plaques can be abortive infection (in which bacterial death leads to phage inactivation) and bacteriocidal compounds in the phage solution (neither situation was tested for in this study) (66). It is interesting that in all but one case (*M. fortuitum* M6 with Power), the phage clearing on *M. fortuitum* was not accompanied by positive screening turbidity assay results. Positive screening turbidity assays were not typically accompanied by positive screening plaque assays, except in the case of the *M. smegmatis* pairings. These tests may reflect that different types of phage-host interactions are occurring in different pairings, or that some phages adsorb to a potential host better in liquid medium, while others adsorb better on a semisolid medium such as agar. Supporting the latter idea, *M. fortuitum* pairings tended to produce results on the screening plaque assay but not the screening turbidity assay, while MuMC pairings tended to do the reverse; the same phage often produced different results in different species (Cucurbita, Schmidt, Shaobing, and Rando14) (Table 16). Different hosts may require different testing conditions.

Only three phage-host pairings produced positive results on the dilution turbidity assay. Positive results on the dilution turbidity assay indicate that these pairings have the potential to represent authentic phage-host interactions and should be evaluated further to illuminate the nature of the potential interaction. The *M. chelonae* M3/Hawkeye (Fig. 16A) and *M. marinum* ATCC927/Noely (Fig. 16C) pairings only produced statistically significant results in the final timepoints, as bacterial growth was transitioning to the stationary phase. Irregular fluctuations of bacterial growth at the point of nutrient limitation may explain these results, or the phage may have exploited the greater concentration of hosts or the decreased health of the hosts in the nutrient-

limited environment. Conversely, the *M. marinum* M/Cucurbita pairing produced a difference between the experimental and control groups on day two, which persisted to the end of the experiment (Fig. 16B). This result could represent that the phage population was increasing as the host population increased. As further evidence for productive phage infection, in the *M. marinum* ATCC927 pairings (Cucurbita in Fig. 16B and Noely in Fig. 16C), the OD_{595} values on a given day tend to sort by phage dilution, with higher phage concentrations having the lower OD_{595} values and lower phage concentrations having the higher OD_{595} values. Interestingly, none of the three host-phage pairings that produced positive results on the dilution turbidity assay produced any plaque assay results during these studies.

The method of phage selection did not appear to affect the results of the host range study (Table 15, Table 16, and Fig. 16). Some of the phages selected based on HostPhinder results (Cucurbita and Shaobing) produced many positive results on the screening, but only Cucurbita had a positive result on the dilution turbidity assay. In fact, the only phages that produced positive results on the dilution turbidity assay were Noely, Cucurbita, and Hawkeye, representing the three methods of phage selection used in this study. A larger sample size and a larger number of selection methods may be necessary to determine an optimal method of host selection for host range studies. HostPhinder may have produced more successful results if the species used in this study had been represented in their database of hosts. Mendokysei was predicted by HostPhinder to have a broad host range comprising five genera but did not have a higher number of positive results on the screening compared to the other phages.

However, Cucurbita was also predicted by HostPhinder to have a large host range (four genera) and did have a higher-than average number of positive results.

The lifestyle of the phage may have had an effect on the host range results, although the sample size was too small to be confident in this inference (Table 16, Fig. 16, and Table 14). Although the results of the screening were not substantially different for temperate versus obligately lytic phages, two of the three phages that produced positive results on the dilution turbidity assay were obligately lytic, and were the only phages used in this study that were known to be obligately lytic (lifestyle data were taken from phagesdb.org and were based on genetic analysis) (Table 14).

Mycobacteriophages of cluster K and subclusters A2 and A3 are known to have a broad host range comprising both *M. smegmatis* and *M. tuberculosis* (137), which suggested that Power (subcluster A2), Shaobing (cluster K), Omnicron (cluster K), and Rando14 (cluster K) would have more positive outcomes in this host range study (Table 14), However, while all four of these phages produced areas of clearing on *M. fortuitum* at high phage concentrations (Table 16), none of these phages had positive results on the dilution turbidity assay (Fig. 16). While these four phages should be able to infect *M. tuberculosis* based on their cluster assignment, HostPhinder did not designate *M. tuberculosis* as a potential host for Omnicron or Rando14. The ability of a phage to complete a productive lytic cycle in a host requires success at each step of this complex process; learning more about phage-host interaction may be required to enhance the predictive ability of host prediction methods.

Unexpectedly, two of the three phages with positive results on the dilution turbidity assay had been initially isolated on hosts outside of the genus *Mycobacterium*

(Table 14 and Fig. 16). Noely was isolated on *Arthrobacter globiformis* B-2979 and Cucurbita was isolated on *Gordonia terrae* 3612 (phagesdb.org). All three genera are in the same order (*Actinomycetales*), but *Mycobacterium* and *Gordonia* are also in the same suborder (*Corynebacterineae*) and both have mycolic acid-containing cell walls (3). *Gordonia* phages, *Arthrobacter* phages, and *Mycobacterium* phages have little genetic similarity but some of these phages have shared hosts (138).

In the present research, plaque assays and turbidity assays in 96-well plates provided markedly different results for host range analysis. Phage-host pairings tended to produce results on one method or the other, but rarely both except in the case of the *M. smegmatis* pairings. Strong phage-host interactions may be visible by multiple methods, while weak interactions may be more greatly influenced by the preferred conditions of the phage. Phages that may infect MuMC, *M. chelonae*, and *M. fortuitum* were identified. Understanding the diverse requirements of phages will enhance the ability to detect a range of phage-host interactions in a variety of hosts.

CHAPTER 6

CONCLUSION

The enormous diversity of phages is becoming increasingly apparent; microscopy, metagenomics, and culture techniques each contribute to forming a clearer view of phages and their interactions with their hosts (22). Special conditions have been identified that permit or enhance the propagation of certain phages (34, 40-42) and it is likely that other phages have requirements that have not yet been discovered. The aims of the present research were to attempt to isolate phages of MuMC, *M. fortuitum*, and *M. chelonae*, to develop new phage culture techniques to aid in these attempts, and to investigate the history of phage infection via bioinformatic analysis.

Turbidity assays in multi-well plates were shown to be efficient, inexpensive, and adaptable to numerous objectives. These techniques could be used in conjunction with plaque assays to determine phage-host interactions in different media types, or as an alternative to plaque assays if the turbidity assay is determined to be more successful for certain phage types.

Attempting many phage enrichment and prophage induction techniques with MuMC, *M. fortuitum*, and *M. chelonae* demonstrated the difficulty some host species present with regards to phage isolation. Multi-well turbidity assays sometimes gave an indication of possible phage-host interaction even when plaque assays produced no visible results.

Many prophages were identified in MuMC genomes, demonstrating that phages have infected these strains in the past. The range of prophage sizes, in addition to the

finding that only the most closely related strains could have vertically inherited prophages, suggests that phage infection is continuously occurring in MuMC and the scarcity of isolated phages is not due to a lack of compatible phages. Many of the prophage genes were almost identical to the genes of known phages of other bacterial species, indicating that MuMC phages are not completely unique. Therefore, there are likely phages present in the environment that can infect MuMC and are not considerably different from phages of related species.

The host range study indicated that some of the phages may have some level of interaction with the tested bacterial strains, but the nature of this interaction is still unknown. Interestingly, these interactions could often be observed by one of the tested methods but not by both (except in the case of *M. smegmatis* pairings and the Power-*M. fortuitum* M6 pairing). Strong interactions resulting in complete, productive lytic cycles may be less influenced by detection method than incomplete infections. Continued optimization of the testing conditions to the host species may improve detection of phage infection and other interactions.

In summary, multi-well plate methods were developed and then used to identify phages that may infect or interact with MuMC, *M. fortuitum*, and *M. chelonae*, and prophages and their characteristics and phylogenies were evaluated in MuMC. The results of these experiments indicate that there are phages capable of infecting these bacterial species, but the isolation of these phages is hindered for unknown reasons. A phage may adapt its mechanism of host adsorption to the environmental conditions in which it is usually found. Turbidity assays sometimes allow the observation of phage-host interactions when plaque assays do not, especially in the case of MuMC. Perhaps

phages of aquatic hosts are better adapted to a liquid medium. The continued optimization of phage isolation techniques will improve the number and diversity of cultured phages, benefiting medicine, agriculture, biotechnology, evolutionary research, and more.

REFERENCES

1. Gauthier DT. 2015. Bacterial zoonoses of fishes: a review and appraisal of evidence for linkages between fish and human infections. Vet. J. 203:27-35. doi:10.1016/j.tvjl.2014.10.028

2. Bönicke R. 1969. Lysogeny among mycobacteria. R. Folia Microbiol. 14:297. doi:10.1007/BF02872695

3. Twort F. 1915. An investigation on the nature of ultra-microscopic viruses. Lancet 186:4814.

4. D'Herelle F. 1917. Sur un microbe invisible antagoniste des bacillus dysentérique. Acad Sci Paris 165:373–5.

5. Salmond GP, Fineran PC. 2015. A century of the phage: past, present and future. Nat. Rev. Microbiol. 13:777-786. doi:10.1038/nrmicro3564

6. Keen EC. 2015. A century of phage research: bacteriophages and the shaping of modern biology. Bioessays 37:6-9. doi:10.1002/bies.201400152

7. Simmonds P, Aiewsakun P. 2018. Virus classification - where do you draw the line? Arch. Virol. 163:2037–2046. doi:10.1007/s00705-018-3938-z

8. Dion MB, Oechslin F, Moineau S. 2020. Phage diversity, genomics and phylogeny. Nat. Rev. Microbiol. 18:125-138. doi:10.1038/s41579-019-0311-5

9. Nobrega FL, Vlot M, de Jonge PA, Dreesens LL, Beaumont HJE, Lavigne R, Dutilh BE, Brouns SJJ. 2018. Targeting mechanisms of tailed bacteriophages. Nat. Rev. Microbiol. 16:760–773. doi:10.1038/s41579-018-0070-8

10. Hatfull GF, Hendrix RW. 2011. Bacteriophages and their genomes. Curr. Opin. Virol. 1:298–303. doi:10.1016/j.coviro.2011.06.009

11. Harada LK, Silva EC, Campos WF, Del Fiol FS, Vila M, Dąbrowska K, Krylov VN, Balcao VM. 2018. Biotechnological applications of bacteriophages: state of the art. Microbiol. Res. 212-213:38-58. doi:10.1016/j.micres.2018.04.007

12. Petrovski S, Dyson ZA, Seviour RJ, Tillett D. 2011. Small but sufficient: the *Rhodococcus* phage RRH1 has the smallest known *Siphoviridae* genome at 14.2 kilobases. J. Virol. 86:358-363. doi:10.1128/JVI.05460-11

13. Al-Shayeb B, Sachdeva R, Chen LX, Ward F, Munk P, Devoto A, Castelle CJ, Olm MR, Bouma-Gregson K, Amano Y, He C, Méheust R, Brooks B, Thomas A, Lavy A, Matheus-Carnevali P, Sun C, Goltsman DSA, Borton MA, Sharrar A, Jaffe AL, Nelson TC, Kantor R, Keren R, Lane KR, Farag IF, Lei S, Finstad K, Amundson R, Anantharaman K, Zhou J, Probst AJ, Power ME, Tringe SG, Li WJ, Wrighton K, Harrison S, Morowitz M, Relman DA, Doudna JA, Lehours AC, Warren L, Cate JHD, Santini JM, Banfield JF. 2020. Clades of huge phages from across Earth's ecosystems. Nature 578:425–431. doi:10.1038/s41586-020-2007-4

14. Breitbart M, Bonnain C, Malki K, Sawaya NA. 2018. Phage puppet masters of the marine microbial realm. Nat. Microbiol. 3:754–766. doi:10.1038/s41564-018-0166-y

15. Dowah ASA, Clokie MRJ. 2018. Review of the nature, diversity and structure of bacteriophage receptor binding proteins that target Gram-positive bacteria. Biophys. Rev. 10:535–542. doi:10.1007/s12551-017-0382-3

16. Davies EV, Winstanley C, Fothergill JL, James CE. 2016. The role of temperate bacteriophages in bacterial infection. FEMS Microbiol. Lett. 363:fnw015. doi:10.1093/femsle/fnw015

17. Fernández L, Rodríguez A, García P. 2018. Phage or foe: an insight into the impact of viral predation on microbial communities. ISME J. 12:1171–1179. https://doi.org/10.1038/s41396-018-0049-5

18. Feiner R, Argov T, Rabinovich L, Sigal N, Brorvok I, Herskovits AA. 2015. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. Nat. Rev. Microbiol. 13:641–650. doi: 10.1038/nrmicro3527.

19. Ofir G, Sorek R. 2018. Contemporary phage biology: from classic models to new insights. Cell 172:1260-1270. doi:10.1016/j.cell.2017.10.045

20. Olszak T, Latka A, Roszniowski B, Valvano MA, Drulis-Kawa Z. 2017. Phage life cycles behind bacterial biodiversity. Curr. Med. Chem. 24:3987–4001. doi:10.2174/0929867324666170413100136

21. Bordenstein S, Bordenstein S. 2016. Eukaryotic association module in phage WO genomes from *Wolbachia*. Nat. Commun. 7:13155. doi:10.1038/ncomms13155

22. Clokie MR, Millard AD, Letarov AV, Heaphy S. 2011. Phages in nature. Bacteriophage 1:31-45. doi:10.4161/bact.1.1.14942

23. Perez-Sepulveda B, Redgwell T, Rihtman B, Pitt F, Scanlan DJ, Millard A. 2016. Marine phage genomics: the tip of the iceberg. FEMS Microbiol. Lett. 363:fnw158. doi:10.1093/femsle/fnw158

24. Hanauer DI, Graham MJ, Betancur L, Bobrownicki A, Cresawn SG, Garlena RA, Jacobs-Sera D, Kaufmann N, Pope WH, Russell DA, Jacobs WR, Sivanathan V, Asai

DJ, Hatfull GF. 2017. An inclusive Research Education Community (iREC): impact of the SEA-PHAGES program on research outcomes and student learning. Proc. Natl. Acad. Sci. U. S. A. 114:13531-13536. doi:10.1073/pnas.1718188115

25. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage Discovery Guide. Howard Hughes Medical Institute. Retrieved from https://seaphagesphagediscoveryguide.helpdocsonline.com/home

26. National Center for Biotechnology Information (NCBI) [Internet]. 1988. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information [cited 2020 Jun 22]. https://www.ncbi.nlm.nih.gov/

27. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs Jr WR, Hendrix RW, Lawrence JG, Hatfull GF. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife 4:e06416. doi:10.7554/eLife.06416

28. Lauster D, Klenk S, Ludwig K, Nojoumi S, Behren S, Adam L, Stadtmüller M, Saenger S, Zimmler S, Hönzke K, Yao L, Hoffmann U, Bardua M, Hamann A, Witzenrath M, Sander LE, Wolff T, Hocke AC, Hippenstiel S, De Carlo S, Neudecker J, Osterrieder K, Budisa N, Netz RR, Böttcher C, Liese S, Herrmann A, Hackenberger CPR. 2020. Phage capsid nanoparticles with defined ligand arrangement block influenza virus entry. Nat. Nanotechnol. 15:373–379. doi:10.1038/s41565-020-0660-2

29. Parasion S, Kwiatek M, Gryko R, Mizak L, Malm A. 2014. Bacteriophages as an alternative strategy for fighting biofilm development. Pol. J. Microbiol. 63:137–145. doi:10.33073/pjm-2014-019

30. Kiefer B, Dahl JL. 2015. Disruption of *Mycobacterium smegmatis* biofilms using bacteriophages alone or in combination with mechanical stress. Adv. Microbiol. 5:699-710. doi:10.4236/aim.2015.510073.

31. Nilsson AS. 2014. Phage therapy--constraints and possibilities. Ups. J. Med. Sci. 119:192-198. doi: 10.3109/03009734.2014.902878

32. McCallin S, Sacher JC, Zheng J, Chan BK. 2019. Current state of compassionate phage therapy. Viruses 11:343. doi:10.3390/v11040343

33. Maciejewska B, Olszak T, Drulis-Kawa Z. 2018. Applications of bacteriophages versus phage enzymes to combat and cure bacterial infections: an ambitious and also a realistic application? Appl. Microbiol. Biotechnol. 102:2563–2581. doi:10.1007/s00253-018-8811-1

34. Serwer P, Hayes SJ, Thomas JA, Hardies SC. 2007. Propagating the missing bacteriophages: a large bacteriophage in a new class. Virol. J 4:21. doi: 10.1186/1743-422X-4-21

35. Hargreaves KR, Clokie MR. 2014. *Clostridium difficile* phages: still difficult? Front. Microbiol. 5:184. doi:10.3389/fmicb.2014.00184

36. Mahony J, Lugli GA, van Sinderen D, Ventura M. 2018. Impact of gut-associated bifidobacteria and their phages on health: two sides of the same coin? Appl. Microbiol. Biotechnol. 102:2091-2099. doi:10.1007/s00253-018-8795-x

37. Mavrich TN, Casey E, Oliveira J, Bottacini F, James K, Franz CMAP, Lugli GA, Neve H, Ventura M, Hatfull GF, Mahony J, van Sinderan D. 2018. Characterization and induction of prophages in human gut-associated *Bifidobacterium* hosts. Sci. Rep. 8:12772. doi:10.1038/s41598-018-31181-3

38. Kim JG, Kim SJ, Cvirkaite-Krupovic V, Yu WJ, Gwak JH, López-Pérez M, Rodriguez, Valera F, Krupovic M, Cho JC, Rhee SK. 2019. Spindle-shaped viruses infect marine ammonia-oxidizing thaumarchaea. Proc. Natl. Acad. Sci. U. S. A. 116:15645-15650. doi:10.1073/pnas.1905682116

39. Abedon ST, Yin J. 2009. Bacteriophage plaques: theory and analysis. Methods Mol. Biol. 501:161-174. doi:10.1007/978-1-60327-164-6_17

40. Lillehaug D. 1997. An improved plaque assay for poor plaque-producing temperate lactococcal bacteriophages. J. Appl. Microbiol. 83:85-90. doi: 10.1046/j.1365-2672.1997.00193.x

41. Łoś JM, Golec P, Węgrzyn G, Węgrzyn A, Łoś M. 2008. Simple method for plating *Escherichia coli* bacteriophages forming very small plaques or no plaques under standard conditions. Appl. Environ. Microbiol. 74:5113-5120. doi:10.1128/AEM.00306-08

42. Almeida GMF, Laanto E, Ashrafi R, Sundberg LR. 2019. Bacteriophage adherence to mucus mediates preventive protection against pathogenic bacteria. mBio 10:e01984-19. doi:10.1128/mBio.01984-19

43. Silva YJ, Costa L, Pereira C, Cunha A, Calado R, Gomes NC, Almeida A. 2014. Influence of environmental variables in the efficiency of phage therapy in aquaculture. Microb. Biotechnol. 7:401-413. doi:10.1111/1751-7915.12090

44. Turenne CY. 2019. Nontuberculous mycobacteria: Insights on taxonomy and evolution. Infect. Genet. Evol. 72:159-168. doi:10.1016/j.meegid.2019.01.017

45. Runyon EH. 1959. Anonymous mycobacteria in pulmonary disease. Med. Clin. N. Am. 43:273-290. doi:10.1016/s0025-7125(16)34193-1

46. Chiner-Oms Á, Comas I. 2019. Large genomics datasets shed light on the evolution of the *Mycobacterium tuberculosis* complex. Infect. Genet. Evol. 72:10-15. doi:10.1016/j.meegid.2019.02.028.

47. Azimi T, Mosadegh M, Nasiri MJ, Sabour S, Karimaei S, Nasser A. 2019. Phage therapy as a renewed therapeutic approach to mycobacterial infections: a comprehensive review. Infect. Drug Resist. 12:2943–2959. doi:10.2147/IDR.S218638

48. Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, Gilmour KC, Soothill J, Jacobs-Sera D, Schooley RT, Hatfull GF, Spencer H. 2019. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. *Nat. Med.* 25:730-733. doi:10.1038/s41591-019-0437-z

49. Puiu M, Julius C. 2019. Bacteriophage gene products as potential antimicrobials against tuberculosis. Biochem. Soc. Trans. 47:847-860. doi:10.1042/BST20180506.

50. Fu X, Ding M, Zhang N, Li J. 2015. Mycobacteriophages: an important tool for the diagnosis of *Mycobacterium tuberculosis* (review). Mol. Med. Rep. 12:13-19. doi:10.3892/mmr.2015.3440

51. Honda JR, Virdi R, Chan ED. 2018. Global environmental nontuberculous mycobacteria and their contemporaneous man-made and natural niches. Front. Microbiol. 9:2029. doi:10.3389/fmicb.2018.02029

52. Falkinham JO. 2016. Current epidemiologic trends of the nontuberculous mycobacteria (NTM). Curr. Environ. Health. Rep 3:161-167. doi:10.1007/s40572-016-0086-z

53. Juhasz SE, Bönicke R. 1965. Possible classification of rapidly growing mycobacteria on the basis of their phage susceptibility. Can. J. Microbiol. 11:235-241. doi:10.1139/m65-030.

54. Rybniker J, Kramme S, Small PL. 2006. Host range of 14 mycobacteriophages in *Mycobacterium ulcerans* and seven other mycobacteria including *Mycobacterium tuberculosis*--application for identification and susceptibility testing. J. Med. Microbiol. 55:37-42. doi:10.1099/jmm.0.46238-0

55. Jeon M, Kim J, Jacobs W, Hartman T, Engel H. 2007. Identification of mycobacteriophages that infect nontuberculous mycobacteria associated with post-LASIK keratitis [Meeting abstract]. Invest. Ophthalmol. Vis. Sci. 48:2662.

56. Grange JM, Bird RG. 1975. The nature and incidence of lysogeny in *Mycobacterium fortuitum*. J. Med. Microbiol. 8:215-223. doi:10.1099/00222615-8-2-215

57. Hyman P. 2019. Phages for phage therapy: isolation, characterization, and host range breadth. Pharmaceuticals (Basel) 12:35. doi:10.3390/ph12010035

58. Branston S, Stanley E, Keshavarz-Moore E, Ward J. 2011. Precipitation of filamentous bacteriophages for their selective recovery in primary purification. Biotechnol. Prog. 28:129–136. doi:10.1002/btpr.705

59. Hall JP, Harrison E, Brockhurst MA. 2013. Viral host-adaptation: insights from evolution experiments with phages. Curr. Opin. Virol. 3:572–577. doi:10.1016/j.coviro.2013.07.001

60. Burrowes BH, Molineux IJ, Fralick JA. 2019. Directed in vitro evolution of therapeutic bacteriophages: the Appelmans protocol. Viruses 11:241. doi:10.3390/v11030241\

61. Antibody Design Labs. 2015. Small scale preparation of bacteriophage by PEG precipitation. Antibody Design Labs Protocols. Accessed February 23, 2020. http://www.abdesignlabs.com/technical-resources/

62. Pettersson BM, Das S, Behra PR, Jordan HR, Ramesh M, Mallick A, Root KM, Cheramie MN, de la Cruz Melara I, Small PLC, Dasgupta S, Ennis DG, Kirsebom LA. 2015. Comparative sigma factor-mRNA levels in *Mycobacterium marinum* under stress conditions and during host infection. PLoS One 10:e0139823. doi:10.1371/journal.pone.0139823

63. Ho CH, Stanton-Cook M, Beatson SA, Bansal N, Turner MS. 2016. Stability of active prophages in industrial *Lactococcus* lactis strains in the presence of heat, acid, osmotic, oxidative and antibiotic stressors. Int. J. Food Microbiol. 220:26-32. doi:10.1016/j.ijfoodmicro.2015.12.012

64. Fusco S, Aulitto M, Bartolucci S, Contursi P. 2014. A standardized protocol for the UV induction of *Sulfolobus* spindle-shaped virus 1. Extremophiles 19:539–546. doi:10.1007/s00792-014-0717-y

65. IBM Corp. Released 2016. IBM SPSS statistics for Windows, version 24.0. Armonk, NY: IBM Corp.

66. Abedon ST. 2011. Lysis from without. Bacteriophage 1:46-49, doi:10.4161/bact.1.1.13980

67. Canchaya C, Proux C, Fournous G, Bruttin A, Brüssow H. 2003. Prophage genomics. Microbiol. Mol. Biol. Rev 67:238-276. doi:10.1128/mmbr.67.2.238-276.2003

68. Doss J, Culbertson K, Hahn D, Camacho J, Barekzi N. 2017. A review of phage therapy against bacterial pathogens of aquatic and terrestrial organisms. Viruses, 9:50. doi:10.3390/v9030050

69. Mattila S, Ruotsalainen P, Jalasvuori M. 2015. On-demand isolation of bacteriophages against drug-resistant bacteria for personalized phage therapy. Front. Microbiol. 6:1271. doi:10.3389/fmicb.2015.01271

70. Giraffa G, Rossetti L. 2004. Monitoring of the bacterial composition of dairy starter cultures by RAPD-PCR. FEMS Microbiol. Lett. 237:133-138. doi:10.1016/j.femsle.2004.06.022

71. Appelmans R. 1921. Le dosage du bacteriophages. Compt. Rend. Soc. Biol. 85:1098–1099.

72. D'Hérelle F, Twort FW, Bordet J, Gratia A, Ledingham JCG, McLeod JW. 1922. Discussion on the bacteriophage (bacteriolysin). BMJ 2:289-299.

73. Scotti PD. 1977. End-point dilution and plaque assay methods for titration of cricket paralysis virus in cultured *Drosophila* cells. J. Gen. Virol. 35:393-396. doi:10.1099/0022-1317-35-2-393.

74. Bonilla N, Rojas MI, Netto Flores Cruz G, Hung SH, Rohwer F, Barr JJ. 2016. Phage on tap-a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. PeerJ 4:e2261. doi:10.7717/peerj.2261.

75. Mizuno CM, Rodriguez-Valera F, Kimes NE, Ghai R. 2013. Expanding the marine virosphere using metagenomics. PLoS Genet. 9:e1003987. doi:10.1371/journal.pgen.1003987

76. Bruder K, Malki K, Cooper A, Sible E, Shapiro JW, Watkins SC, Putonti C. 2016. Freshwater metaviromics and bacteriophages: a current assessment of the state of the art in relation to bioinformatic challenges. Evol. Bioinform. Online 12:25-33. doi:10.4137/EBO.S38549

77. Jäckel C, Hammerl JA, Rau J, Hertwig S. 2017. A multiplex real-time PCR for the detection and differentiation of *Campylobacter* phages. PLoS One 12:e0190240. doi:10.1371/journal.pone.0190240

78. Abbineni G, Safiejko-Mroczka B, Mao C. 2010. Development of an optimized protocol for studying the interaction of filamentous bacteriophage with mammalian cells by fluorescence microscopy. Microsc. Res. Tech. 73:548-554. doi:10.1002/jemt.20793

79. Lu Z, Breidt F, Plengvidhya V, Fleming HP. 2003. Bacteriophage ecology in commercial sauerkraut fermentations. Appl. Environ. Microbiol. 69:3192-3202. doi:10.1128/AEM.69.6.3192-3202.2003

80. McDonald JE, Smith DL, Fogg PC, McCarthy AJ, Allison HE. 2010. High-throughput method for rapid induction of prophages from lysogens and its application in the study of Shiga toxin-encoding *Escherichia coli* strains. Appl. Environ. Microbiol. 76:2360-2365. doi:10.1128/AEM.02923-09

81. Heselpoth RD, Nelson DC. 2012. A new screening method for the directed evolution of thermostable bacteriolytic enzymes. J. Vis. Exp. 69:4216. doi:10.3791/4216

82. Pérez-Gamarra S, Hattara L, Batra G, Saviranta P, Lamminmaki U. 2017. Array-in-well binding assay for multiparameter screening of phage displayed antibodies. Methods 116:43-50. doi:10.1016/j.ymeth.2016.12.004

83. Fischer S, Kittler S, Klein G, Glunder G. 2013. Microplate-test for the rapid determination of bacteriophage-susceptibility of *Campylobacter* isolates-development and validation. PLoS One 8:e53899. doi:10.1371/journal.pone.0053899

84. Sarkis GJ, Jacobs Jr WR, Hatfull GF. 1995. L5 luciferase reporter mycobacteriophages: a sensitive tool for the detection and assay of live mycobacteria. Mol. Microbiol. 15:1055-1067. doi:10.1111/j.1365-2958.1995.tb02281.x

85. Turpin PE, Maycroft KA, Bedford J, Rowlands CL, Wellington EMH. 1993. A rapid luminescent-phage based MPN method for the enumeration of *Salmonella typhimurium* in environmental samples. Lett. Appl. Microbiol. 16:24-27. doi:10.1111/j.1472-765X.1993.tb01364.x

86. McLaughlin MR. 2007. Simple colorimetric microplate test of phage lysis in *Salmonella enterica*. J. Microbiol. Methods 69:394-398. doi:10.1016/j.mimet.2007.01.006

87. Cooper CJ, Denyer SP, Maillard JY. 2011. Rapid and quantitative automated measurement of bacteriophage activity against cystic fibrosis isolates of *Pseudomonas aeruginosa*. J. Appl. Microbiol. 110:631-640. doi:10.1111/j.1365-2672.2010.04928.x

88. Henry M, Biswas B, Vincent L, Mokashi V, Schuch R, Bishop-Lilly KA, Sozhamannan S. 2012. Development of a high throughput assay for indirectly measuring phage growth using the OmniLog(TM) system. Bacteriophage 2:159-167. doi:10.4161/bact.21440

89. Xie Y, Wahab L, Gill JJ. 2018. Development and validation of a microtiter plate-based assay for determination of bacteriophage host range and virulence. Viruses 10:189. doi:10.3390/v10040189

90. Crowley GC, O'Mahony J, Coffey A, Sayers R, Cotter P. 2019. A rapid viability and drug-susceptibility assay utilizing mycobacteriophage as an indicator of drug susceptibilities of anti-TB drugs against *Mycobacterium smegmatis* mc^2155. Int. J. Mycobacteriol. 8:124-131. doi:10.4103/ijmy.ijmy_47_19

91. Filée J, Tétart F, Suttle CA, Krisch HM. 2005. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. Proc. Natl. Acad. Sci. U S A 102:12471-12476. doi:10.1073/pnas.0503404102

92. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784-786. doi:10.1093/bioinformatics/btw711

93. Tortoli E. 2014. Microbiological features and clinical relevance of new species of the genus *Mycobacterium*. Clin. Microbiol. Rev. 27:727-752. doi:10.1128/CMR.00035-14

94. Rivero-Lezcano OM, González-Cortés C, Mirsaeidi M. 2019. The unexplained increase of nontuberculous mycobacteriosis. Int. J. Mycobacteriol 8:1-6. doi:10.4103/ijmy.ijmy_18_19

95. Guo S, Ao Z. 2012. Phage in the diagnosis and treatment of tuberculosis. Front. Biosci. 17:2691-2697. doi:10.2741/4080

96. Rajnovic D, Muñoz-Berbel X, Mas J. 2019. Fast phage detection and quantification: an optical density-based approach. PLoS One 14:e0216292. doi:10.1371/journal.pone.0216292

97. Fortier LC, Sekulovic O. 2013. Importance of prophages to evolution and virulence of bacterial pathogens. Virulence 4:354-365. doi:10.4161/viru.24498

98. Argov T, Azulay G, Pasechnek A, Stadnyuk O, Ran-Sapir S, Borovok I, Sigal N, Herskovits AA. 2017. Temperate bacteriophages as regulators of host behavior. Curr. Opin. Microbiol. 38:81-87. doi:10.1016/j.mib.2017.05.002

99. Hurwitz, BL., Ponsero A, Thornton J, U'Ren JM. 2017. Phage hunters: computational strategies for finding phages in large-scale 'omics datasets. Virus Res. 15:110-115. https://doi.org/10.1016/j.virusres.2017.10.019

100. Casjens S. 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. 49:277-300. doi: 10.1046/j.1365-2958.2003.03580.x

101. Arkhipova, IR. 2017. Using bioinformatic and phylogenetic approaches to classify transposable elements and understand their complex evolutionary histories. Mob. DNA 8:19. doi:10.1186/s13100-017-0103-2

102. Bobay LM, Touchon M, Rocha EP. 2014. Pervasive domestication of defective prophages by bacteria. Proc. Natl. Acad. Sci. U. S. A. 111:12127–12132. doi:10.1073/pnas.1405336111

103. Shaaban S, Cowley LA., McAteer SP, Jenkins C, Dallman TJ, Bono JL, Gally DL. 2016. Evolution of a zoonotic pathogen: investigating prophage diversity in enterohaemorrhagic *Escherichia coli* O157 by long-read sequencing. Microb. Genom. 2:e000096. doi:10.1099/mgen.0.000096

104. Fu T, Fan X, Long Q, Deng W, Song J, Huang E. 2017. Comparative analysis of prophages in *Streptococcus mutans* genomes. *PeerJ 5*:e4057. doi:10.7717/peerj.4057

105. Casjens SR, Grose JH. 2016. Contributions of P2- and P22-like prophages to understanding the enormous diversity and abundance of tailed bacteriophages. Virology 496:255–276. doi:10.1016/j.virol.2016.05.022

106. Vale FF, Nunes A, Oleastro M, Gomes J, Sampaio DA, Rocha R, Vítor JMB, Engstrand L, Pascoe B, Berthenet E, Sheppard SK, Hitchings MD, Mégraud F, Vadivelu J, Lehours P. 2017. Genomic structure and insertion sites of *Helicobacter pylori* prophages from various geographical origins. Sci. Rep. 7:42471. doi:10.1038/srep42471

107. Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc. Natl. Acad. Sci. U. S. A. 96:2192-2197. doi:10.1073/pnas.96.5.2192

108. Fan X, Xie L, Li W, Xie J. 2014. Prophage-like elements present in *Mycobacterium* genomes. BMC Genomics 15:243. doi:10.1186/1471-2164-15-243

109. Fan X, Abd Alla AA, Xie J. 2016. Distribution and function of prophage phiRv1 and phiRv2 among *Mycobacterium tuberculosis* complex. J. Biomol. Struct. Dyn. 34:233-238. doi:10.1080/07391102.2015.1022602

110. Sassi M, Gouret P, Chabrol O, Pontarotti P, Drancourt M. 2014. Mycobacteriophage-drived diversification *of Mycobacterium abscessus*. *Biol. Direct* 9:19. doi:10.1186/1745-6150-9-19

111. Voronina OL, Kunda MS, Aksenova EI, Semenov AN, Ryzhova NN, Lunin VG, Gintsburg AL. 2016. Mosaic structure of *Mycobacterium bovis* BCG genomes as a representation of phage sequences' mobility. BMC Genomics 17:1009. doi:10.1186/s12864-016-3355-1

112. Bouam A, Armstrong N, Levasseur A, Drancourt M. 2018. *Mycobacterium terramassiliense*, *Mycobacterium rhizamassiliense* and *Mycobact*

erium numidiamassiliense sp. nov., three new *Mycobacterium simiae* complex species cultured from plant roots. Sci. Rep. 8:9309. doi:10.1038/s41598-018-27629-1

113. Rhodes MW, Kator H, Kaattari I, Gauthier D, Vogelbein WK, Ottinger CA. 2004. Isolation and characterization of mycobacteria from striped bass *Morone saxatilis* from the Chesapeake Bay. Dis. Aquat. Org. 61:41-51. doi:10.3354/dao061041

114. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. 2000. A whole-genome assembly of *Drosophila*. Science 287:2196-2204. doi: 10.1126/science.287.5461.2196.

115. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-2069. doi: 10.1093/bioinformatics/btu153

116. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649. doi:10.1093/bioinformatics/bts199

117. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44:16-21. doi:10.1093/nar/gkw387

118. Stinear TP, Seemann T, Harrison PF, Jenkin GA, Davies JK, Johnson PDR, Abdellah Z, Arrowsmith C, Chillingworth T, Churcher C, Clarke K, Cronin A, Davis

P, Goodhead I, Holroyd N, Jagels K, Lord A, Moule S, Mungall K, Norbertczak H, Quail MA, Rabbinowitsch E, Walker D, White B, Whitehead S, Small PLC, Brosch R, Ramakrishnan L, Fischbach MA, Parkhill J, Cole ST. 2008. Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. Genome Res. 18:729–741. doi:10.1101/gr.075069.107

119. Darling ACE, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14:1394-1403. doi: 10.1101/gr.2289704

120. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410. doi:10.1016/S0022-2836(05)80360-2

121. Huson DH. 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics 14:68-73. doi:10.1093/bioinformatics/14.1.68

122. Huson, DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254-267. doi:10.1093/molbev/msj030

123. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797. doi:10.1093/nar/gkh340

124. Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704. doi:10.1080/10635150390235520

125. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9:772. doi:10.1038/nmeth.2109

126. Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. Nucleic Acids Res. 35:52-57. doi:10.1093/nar/gkm360

127. Das S, Pettersson B, Behra P, Mallick A, Cheramie M, Ramesh M, Shirreff L, DuCote T, Dasgupta S, Ennis DG, Kirsebom LA. 2018. Extensive genomic diversity among *Mycobacterium marinum* strains revealed by whole genome sequencing. Sci. Rep. 8:12040. doi:10.1038/s41598-018-30152-y

128. van der Sar AM, Abdallah AM, Sparrius M, Reinders E, Vandenbroucke-Grauls CM, Bitter W. 2004. *Mycobacterium marinum* strains can be divided into two distinct types based on genetic diversity and virulence. Infect. Immun. 72:6306–6312. doi:10.1128/IAI.72.11.6306-6312.2004

129. Williams, KP. 2002. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucleic Acids Res. 30:866–875. doi:10.1093/nar/30.4.866

130. Briner AE, Lugli GA, Milani C, Duranti S, Turroni F, Gueimonde M, Margolles A, van Sinderen D, Ventura M, & Barrangou R. 2015. Occurrence and diversity of CRISPR-Cas systems in the genus *Bifidobacterium*. PLoS ONE 10:e0133661. doi:10.1371/journal.pone.0133661

131. Nozawa T, Furukawa N, Aikawa C, Watanabe T, Haobam B, Kurokawa K, Maruyama F, Nakagawa I. 2011. CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. PLoS ONE 6:e19543. doi: 10.1371/journal.pone.0019543

132. Hatoum-Aslan A, Marraffini LA. 2014. Impact of CRISPR immunity on the emergence and virulence of bacterial pathogens. Curr. Opin. Microbiol. 17:82-90. doi:10.1016/j.mib.2013.12.001

133. O'Meara D, Nunney L. 2019. A phylogenetic test of the role of CRISPR-Cas in limiting plasmid acquisition and prophage integration in bacteria. Plasmid 104:102418. doi:10.1016/j.plasmid.2019.102418

134. Rezaei Javan R, Ramos-Sevillano E, Akter A, Brown J, Brueggeman AB. 2019. Prophages and satellite prophages are widespread in *Streptococcus* and may play a role in pneumococcal pathogenesis. Nat. Commun. 10:4852. doi:10.1038/s41467-019-12825-y

135. Timme TL, Brennan PJ. 1984. Induction of bacteriophage from members of the *Mycobacterium avium*, *Mycobacterium intracellulare*, *Mycobacterium scrofulaceum* serocomplex. Microbiology 130:2059–2066. doi:10.1099/00221287-130-8-2059

136. Villarroel J, Kleinheinz KA, Jurtz VI, Zschach H, Lund O, Nielson M, Larsen MV. 2016. HostPhinder: a phage host prediction tool. Viruses 8:116. doi:10.3390/v8050116

137. Jacobs-Sera D, Marinelli LJ, Bowman C, Broussard G, Guerrero C, Boyle M, Petrova Z, Dedrick R, Pope W, Modlin RL, Hendrix RW, Hatfull GF. 2012. On the nature of mycobacteriophage diversity and host preference. Virology 434:187-201. doi:10.1016/j.virol.2012.09.026

138. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Hatfull GF. 2017. Bacteriophages of *Gordonia* spp. display a spectrum of diversity and genetic relationships. mBio 8:e01069-17. doi:10.1128/mBio.01069-17

VITA

Janis H. Doss

Department of Biological Sciences

106 Mills Godwin Building, Old Dominion University, Norfolk, VA 23529

EDUCATION

Current	Doctor of Philosophy, Biomedical Sciences
	Old Dominion University, Norfolk, VA
05/2012	Associate of Applied Science, Medical Laboratory Technology
	Coastal Carolina Community College, Jacksonville, NC
05/2007	Bachelor of Science, Biology, Vertebrate Physiology Option
	Minor in Psychology
	The Pennsylvania State University, University Park, PA

PUBLICATION

Doss J, Culbertson K, Hahn D, Camacho J, Barekzi N. 2017. A Review of Phage Therapy against Bacterial Pathogens of Aquatic and Terrestrial Organisms. *Viruses* 9:50. doi:10.3390/v9030050

SELECTED PRESENTATIONS

08/2019	Evergreen International Phage Biology Meeting, Olympia, WA Poster presentation “Phylogeny and Evolution of Prophages in 49 Genomes of *Mycobacterium marinum*”
06/2017	American Society for Microbiology - Microbe 2017, New Orleans, LA Poster presentation “The Use of 96-well Plates for the Detection of Bacteriophages in Liquid Medium”
06/2016	American Society for Microbiology - Microbe 2016, Boston, MA Poster presentation “Genomic Effects of Host-Specific Specialization in *Mycobacterium shottsii*”