Optimizing high-throughput detection of temperate phages that infect *Enterococcus* spp. and *Staphylococcus* spp.

by

Tia Arnaud

A Thesis submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Microbiology

University of Manitoba

Winnipeg, Manitoba, Canada

ABSTRACT

Bacteriophage (phage) therapy is currently being explored as an alternative treatment strategy to combat infections caused by antibiotic-resistant bacteria. However, considerable emphasis has been placed on utilizing only strictly virulent phages for treatment whereas the use of temperate phages has been largely inadvisable due to their potential to transmit antibiotic resistance genes. Unfortunately, the current process of isolating and deducing the lifecycle of newly-isolated phages is often labour-intensive, costly, and time-consuming, highlighting the need for a more streamlined process. Therefore, the aim of this study was to develop a PCR-based assay to screen out temperate phages that infect *Enterococcus faecalis* and *Staphylococcus aureus* during high-throughput phage isolation. Using the integrase gene in temperate phages as a molecular marker, cluster analysis was performed to identify conserved sequences and design cluster-specific degenerate primers. Multiplex PCRs (mPCRs) were then optimized using a set of short DNA sequences representing the integrase genes, and tested on a panel of plaque picks from phages isolated against *E. faecalis* and *S. aureus*. mPCR results were compared with whole genome and amplicon sequencing to verify the presence integrase genes. Overall, this study revealed that specific primers can be designed to target integrase genes in a simple and rapid mPCR assay. However, host-encoded prophages were detected from contaminating host nucleic acid in plaque picks, resulting in false positives as indicated by the sequencing results. This method may still be adopted – albeit, with careful consideration – by assuming that any same-sized band amplified from both the plaque picks and bacterial hosts are the result of contamination, and are therefore negative results. Ultimately, the findings of this research showed that a PCR screening method that detects homologues present only in virulent phages may be a better alternative.

ACKNOWLEDGEMENTS

I would like to thank my advisors, Dr. Steven Theriault and Dr. Deborah Court, for their unwavering support, guidance, and feedback over the course of this degree. The encouragement to pursue my interests and freedom to be creative were deeply appreciated. I would also like to express my gratitude to committee members Dr. Aleeza Gerstein and Dr. Georg Hausner for the insightful discussions that helped shape the development of this research. A huge thank you to Dr. Tasia Lightly for making time for every impromptu meeting and helping me navigate this degree – my appreciation cannot be overstated. A special thank you to the staff at Cytophage Technologies Inc. for all the support, understanding, and laughs over the years. I would also like to thank Cytophage Technologies Inc. for the use of their laboratory and funding this research, as well as the Department of Microbiology for helping me grow as a researcher. Lastly, I would like to thank my friends and family for their constant encouragement and patience during this challenging journey.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AR	Antibiotic resistance
BV-BRC	Bacterial and Viral Bioinformatics Resource Center
$CaCl_2$	Calcium chloride
CTI	Cytophage Technologies Inc.
Ent-SA	*Enterococcus* serine cluster A integrases
Ent-SB	*Enterococcus* serine cluster B integrases
Ent-TA	*Enterococcus* tyrosine cluster A integrases
Ent-TB	*Enterococcus* tyrosine cluster B integrases
Ent-TC	*Enterococcus* tyrosine cluster C integrases
HAI	Hospital-acquired infections
LB	Luria-Bertani broth
MCP	Major capsid protein
$MgSO_4 \cdot 7H_2O$	Magnesium sulfate heptahydrate
MOI	Multiplicity of infection
mPCR	Multiplex polymerase chain reaction
MRSA	Methicillin-resistant *Staphylococcus aureus*
MSA	Multiple sequence alignment
NaCl	Sodium chloride
ORF	Open reading frame
PCR	Polymerase chain reaction
PFU	Plaque forming units

PHASTER	PHAge Search Tool Enhance Release
PT	Phage therapy
SM buffer	Saline-magnesium buffer
Sta-SA	*Staphylococcus* serine cluster A integrases
Sta-TA	*Staphylococcus* tyrosine cluster A integrases
Sta-TB	*Staphylococcus* tyrosine cluster B integrases
Sta-TC	*Staphylococcus* tyrosine cluster C integrases
TSA	Tryptic soy agar
TSB	Tryptic soy broth
VRE	Vancomycin-resistant enterococci
WGS	Whole genome sequencing
WHO	World Health Organization

CHAPTER 1.0: LITERATURE REVIEW

1.1 Antibiotic resistance: an overview

For decades, the use of antibiotics has played a crucial role in preventing and treating bacterial infections. This has significantly improved patient outcomes by lowering morbidity and mortality rates, and facilitating advancements in cancer treatment, organ transplantation, and surgical procedures (1). However, the rapid emergence of antibiotic resistant bacteria in recent years has rendered antibiotic therapy increasingly ineffective (2). This has been attributed to several factors including the overuse of antibiotics, incorrectly prescribed antibiotics, extensive use as growth promoters in agriculture, poor sanitation and hygiene practices, and lack of novel antimicrobial discovery (1). Consequently, antibiotic resistance (AR) has now been recognized as a major public health threat resulting in increased severity of infections, mortality rates, risks of complications, and healthcare costs (3).

1.2 Priority pathogens for research and development of new antimicrobials

In response to the accelerating threat of AR, the World Health Organization (WHO) published a list of global priority pathogens in 2017 for research and development of novel antimicrobials, categorizing each pathogen into one of three priority tiers – critical, high, or medium – based on the urgency of need for new antibiotics (4,5). The primary aim of this initiative was to prioritize funding, coordinate research efforts to address relevant public health concerns, and encourage global collaboration to combat AR (5). The pathogens identified on this list were determined based on an evaluation of mortality, incidence, non-fatal health burden, resistance, transmissibility, preventability in healthcare and community settings, treatability, and the antibacterial pipeline (4). Among the list of pathogens are three Gram-positive bacteria: vancomycin-resistant *Enterococcus faecium* (high priority), methicillin-resistant and

vancomycin-intermediate and -resistant *Staphylococcus aureus* (high priority), and penicillin-non-susceptible *Streptococcus pneumoniae* (medium priority) (4). Of these three pathogens, the former two will be the focus of this study given their higher priority.

1.2.1 *Enterococcus* spp.

1.2.1.1 Introduction to *Enterococcus* spp.

Comprised of over 60 species, members of the *Enterococcus* genus are classified as Gram-positive, catalase-negative, non-spore-forming, facultative anaerobes (6,7). Additionally, enterococci have been regarded as unusually resistant to various environmental stressors including temperature, pH, salinity, and desiccation (8). Found ubiquitously throughout various ecological niches, these bacteria exist as commensals within the gastrointestinal tracts of humans and animals, although are also opportunistic pathogens (9,10). Of these species, *E. faecalis* and *E. faecium* account for the majority of enterococcal infections, particularly in healthcare settings as one of the leading causes of global nosocomial infections (6,11).

1.2.1.2 Antibiotic resistance

Difficulties in treating enterococcal infections have emerged due to multidrug resistance. This has been attributed to both intrinsic and acquired mechanisms. In brief, enterococci are resistant to several groups of antibiotics including aminoglycosides, lincosamides, glycopeptides, β-lactams and cephalosporins through naturally developed resistance mechanisms including altered penicillin-binding proteins, production of beta-lactamases, production of aminoglycoside-modifying enzymes, and decreased cellular permeability (2,12,13). This has been further compounded through acquired resistance mechanisms such as spontaneous mutations and horizontal gene transfer of antimicrobial resistance genes by plasmids or transposons (12).

Indeed, a previous study observed that more than 92% of *E. faecium* isolates collected from clinical samples of children were resistant to ampicillin (92.5%), ciprofloxacin (96%), erythromycin (100%), and clindamycin (96%), with a similar resistance profile observed in *E. faecalis* isolates (14).

In response to this multidrug resistance, widespread use of vancomycin – a last-resort antibiotic – has been used as an alternative antibiotic by clinicians since the 1970s (15). However, vancomycin-resistant enterococci (VRE) quickly emerged (15). Recent reports indicated that the prevalence of vancomycin resistance has been observed in 10% and 80% of *E. faecalis* and *E. faecium* isolates, respectively (11). In an effort to combat infections caused by VRE, the use of alternative last-resort antibiotics, such as linezolid, tigecycline, and daptomycin, has also been implemented; although, emerging resistance has been reported (16,17).

1.2.1.3 Clinical significance

As the third most common cause of hospital-acquired infections (HAIs), enterococcal species account for 10% of such cases in the US (18). Of these infections, *E. faecalis* has been recovered from 85-90% of clinical isolates, where as *E. faecium* is responsible for approximately 5-10% (19). Similar proportions have been corroborated by other previous reports as well (20,21).

Infections caused by *E. faecalis* and *E. faecium* vary, but are frequently implicated in urinary tract infections, sepsis, bacteremia, endocarditis, abdomen and biliary tract infections, and infections of catheters and other implanted medical devices (11,22). These infections are most commonly seen after surgery due to burn wounds, leg and pressure ulcers, and infections during diagnostic or therapeutic procedures in the urinary tract (23).

1.2.2 *Staphylococcus* spp.

1.2.2.1 Introduction to *Staphylococcus* spp.

The *Staphylococcus* genus is characterized as Gram-positive, non-sporulating, non-motile, facultative anaerobic bacteria (24). Members of this genus may also be further distinguished in clinical settings as coagulase-negative staphylococci or coagulase-positive staphylococci based on their ability to produce the coagulase enzyme, a key virulence factor that aids in promoting abscess formation and infection persistence (25,26). While more than 50 species and subspecies have been identified, only 18 species have been found in humans (27). Found ubiquitously throughout the environment, staphylococci are also highly adaptable to various ecological niches and environmental stressors. Specifically within humans, these bacteria typically colonize the skin and mucous membranes without causing disease in healthy individuals; although, they are also opportunistic pathogens (28).

1.2.2.2 Antibiotic resistance

Similar to *Enterococcus* spp., intrinsic and acquired resistance also contribute to multidrug resistance in clinically relevant staphylococcal isolates. For example, intrinsic resistance in *S. aureus* has been attributed to reduced membrane permeability, efflux systems, and excessive production of β-lactamases (29). In contrast, acquired resistance mechanisms identified in *S. aureus* include mutations, acquisition of resistance genes, biofilm formation, and persister cells as reviewed by Guo *et al.* 2020 (29).

Of particular public health concern is methicillin-resistant *S. aureus* (MRSA), which has developed resistance to all β-lactam antibiotics such as penicillins, cephalosporins, and carbapenems (30). Earlier reports have also identified resistance against macrolides, aminoglycosides, lincosamides, and fluoroquinolones (31). Additionally, previous estimates

indicated that 20% and 33-55% of *S. aureus* isolates in Europe and US hospitals, respectively, are methicillin-resistant (32). This has led to widespread use of vancomycin as an alternative treatment option for clinicians. Albeit rare, the emergence of vancomycin-intermediate *S. aureus,* which exhibits reduced susceptibility, and vancomycin-resistant *S. aureus*, which exhibits complete resistance, has led to increased rates of treatment failure, prolonged hospital stays, and higher healthcare costs (33). Alternatives to vancomycin have also been explored and include a combination treatment of high-dose daptomycin with gentamicin, rifampicin, linezolid, trimethoprim-sulfamethoxazole, or a β-lactam (34). However, multidrug resistance to these alternatives has been observed. For example, a study of MRSA strains isolated from hospital settings revealed that 45.7% of strains were resistant to clindamycin, 64.7% to ciprofloxacin, 87.3% to co-trimoxazole, 54.3% to erythromycin, 17.3% to gentamicin, 16.8% to netilmicin, and 58.38% to tetracycline (35). Clearly, the continued rise of multidrug resistant strains highlights the need to explore alternative therapeutics.

1.2.2.3 Clinical significance

Although unlikely to cause disease in healthy individuals, *S. aureus* and *S. epidermidis* are the two most clinically important staphylococci in humans. Mainly localized to the nasal cavity, groin, throat and perineum, *S. aureus* colonizes approximately 30% of the human population (36,37). In contrast, *S. epidermidis* can be found in 75-100% of individuals at multiple sites including skin, anterior nares, axillae, groin, conjunctiva, and perineum (37,38).

As the leading cause of various nosocomial infections, *S. aureus* is the most invasive staphylococcal species. Concerningly, methicillin-resistant *S. aureus* strains account for 20-80% of hospital-acquired *S. aureus* infections globally (39). Such infections include endocarditis, bacteremia, toxic shock syndrome, osteomyelitis, pneumonia, and deep tissue abscesses (40). In

contrast, *S. epidermidis* is a leading cause of neonatal sepsis and is also associated with endocarditis, prosthetic joint infections, surgical site infections, and catheter infections (41,42).

1.3 Bacteriophages

1.3.1 Introduction to bacteriophages

Bacteriophages (phages) are obligately intracellular, bacteriolytic viruses that specifically infect and kill their bacterial hosts (43). As the most abundant biological entities in nature with an estimated abundance of 10^{31} viral particles, phages are found ubiquitously throughout diverse environments (44). Broadly, phages are composed of proteinaceous or proteolipid capsids containing genomic DNA or RNA (43). Host ranges may also vary among phages. While some are described as having a narrow host range, limited to infecting only a few strains within the same bacterial species, others may exhibit a broad host range and are capable of infecting multiple species and genera (45,46).

1.3.2 Phage diversity

Extensive diversity amongst phages has been discussed throughout literature, underscoring their wide range of genome types and sizes, morphologies, host ranges, lifecycles, and virulent properties (47,48). For example, approximately 55% of phages identified to date morphologically belong to the *Siphoviridae* family, while 25% and 20% belong to the *Myoviridae* and *Podoviridae* families, respectively (48). Additionally, genomes vary considerably and include single and double-stranded DNA or RNA genomes, as well as sizes ranging from approximately 3.3 kb to 735 kb (48,49). Phage diversity has been attributed to the rapid rate of viral evolution resulting from genetic exchanges between phages, allowing for

genome mosaicism; mutations; and co-evolution with bacterial hosts (50–52). Interestingly, the bacteriophage population structure has previously been reconstructed as a network, rather than a phylogenetic tree, to better reflect the diversity and genetic exchanges that can occur between phages (53).

1.3.3 Phage Lifecycles

1.3.3.1 Lytic cycle

Bacteriophage-mediated infection of a host bacterium may follow one of two major lifecycles: the lytic cycle or the lysogenic cycle. In brief, the lytic cycle is comprised of five steps: adsorption, penetration, replication, packaging, and lysis (54). Infection by the lytic cycle initiates with phage adsorption to the host bacterium by irreversibly binding to a specific host cell receptor followed by injection of the viral genetic material into the host (55). Subsequently, host machinery is subverted in order to produce new viral progeny, or virions, which are assembled and packaged into mature viral particles (54). Increased internal osmotic pressure from the accumulation of pore-forming proteins inserted into the cell wall of the host bacterium results in lysis of the host bacterium, ultimately releasing mature viral progeny that can go on to repeat this cycle of replication by infecting other host bacteria (56). Phages that solely replicate through this lifecycle are said to be strictly virulent.

1.3.3.2 Lysogenic cycle

In contrast to strictly virulent phages, temperate phages can enter the lytic or the lysogenic cycle. Entry into the lysogenic cycle is favoured by several environmental factors such as high virus-to-host ratio, known as the multiplicity of infection (MOI), and host nutrient deprivation (57). Once entry into the lysogenic cycle has been initiated, a site-specific recombinase catalyzes integration of the phage genome into the host chromosome where it is

then maintained in a latent state as a prophage (58). In this latent state, genes required for replication by the lytic cycle are repressed (59). Rather, replication of the prophage occurs synchronously with replication of the host genome during cell division, resulting in vertical transmission of the integrated virus to daughter cells without causing cell lysis (60). However, exposure to various stress conditions such as UV irradiation, high temperature, reactive oxygen species, and antibiotics may induce excision of the prophage from the host chromosome and entry into the lytic cycle, ultimately resulting in host cell lysis (61).

1.3.3.2.1 Temperate phage integrases

Phage integrases are site-specific recombinases that are responsible for mediating integration of the temperate phage genome into the bacterial host chromosome during entry of the lysogenic cycle (58). Phage integrases can be further categorized into the following two evolutionarily distinct families based on the conserved catalytic amino acid within the active site of the enzyme: the tyrosine recombinase family or the serine recombinase family (62). Additional details regarding the characteristics and mechanisms of each recombinase family have been previously reviewed in-depth elsewhere; however, discussion is limited to general biological function and diversity of integrases in this review as it is the focus of this research (58).

1.3.3.2.1.1 Biological Function

While the characteristics and recombination mechanisms between the two recombinase families may differ, the process of integration is broadly accomplished through the recognition of sequences in the phage attachment site, *attP,* and the bacterial attachment site, *attB* (58). Once the attachment sites have been recognized, the catalytic residue in the active site mediates a nucleophilic attack to facilitate DNA strand breakage and resolution (58). This ultimately results

in precise recombination and insertion of the phage genome into the host chromosome – flanked by hybrid sites *attL* and *attR* – where it will persist as a prophage (Figure 1) (58).

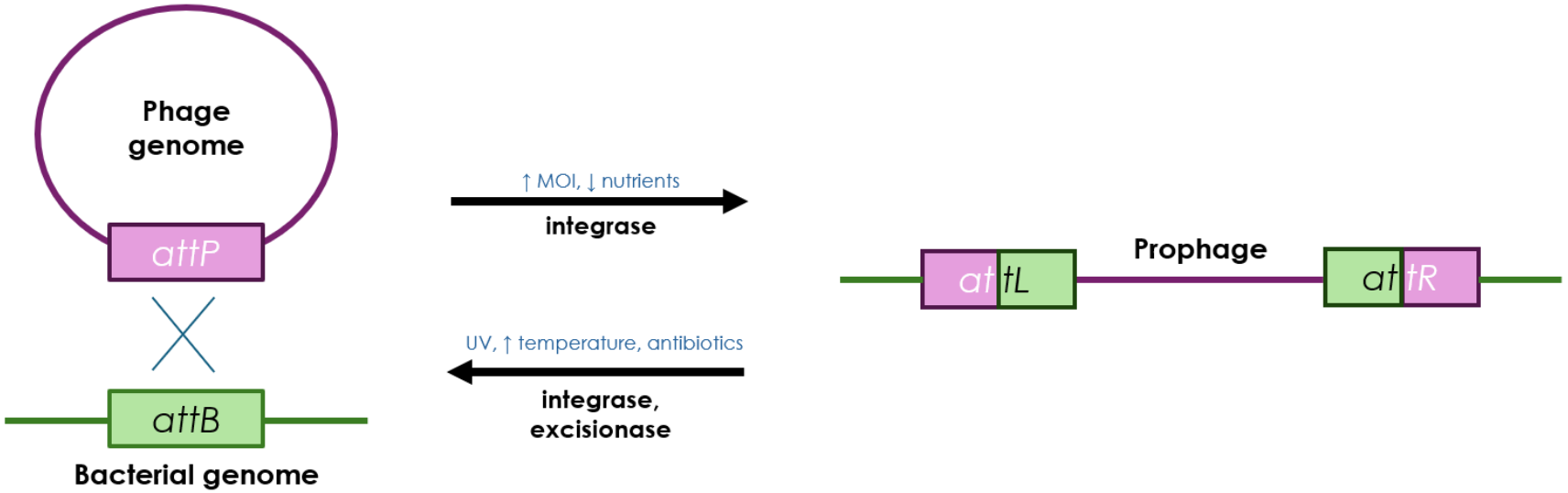

Figure 1. Site-specific integration and excision of a temperate phage into a host chromosome. Infection at a high multiplicity of infection (MOI; virus-to-host ratio) and/or low nutrient availability promotes integration of the phage genome into the bacterial host chromosome by recombination between the phage attachment site (*attP*) and bacterial attachment site (*attB*). This results in the formation of two hybrid attachment sites: *attL* (left attachment site) and *attR* (right attachment site). Excision of the prophage from the host chromosome may be induced by UV irradiation, high temperature, and/or antibiotic treatment. This is mediated by an integrase and excisionase by recombination between the two hybrid sites. See text for reference.

1.3.3.2.1.2 Conservation of integrases across phages

While integrases are conserved in structure and function, little homology between integrase and attachment sequences is present across all temperate phages (58). However, a selective pressure exists to preserve conservation between corresponding *att* sites in order to maintain recognition and recombination between phage and host (58). Additionally, regions of conservation were previously identified in the C-terminal domain; although, the analysis was limited to seven phage site-specific recombinases (63). An examination of taxonomic relationships of enteric phage tyrosine recombinase gene sequences also revealed that while several groups of sequences aligned well to each other, they were overall too divergent to allow for global alignment of all sequences (64).

1.4 Phage therapy

1.4.1 Phage therapy as an alternative to antibiotics

As the rise of multidrug resistance continues to render antibiotic therapy increasingly ineffective, alternative strategies must be explored. One of such alternatives is phage therapy (PT). Conventional PT exploits the bacteriolytic nature of virulent phages to selectively eliminate target pathogenic bacteria that are causing infection (65). In addition to bactericidal properties, some major advantages of phage therapy include auto-dosing, whereby phages will self-replicate at the site of infection; safety and low inherent toxicity to mammalian cells; minimal disruptions to normal microbiota; ubiquity and availability for isolation; and potential for modifiability to personalize treatment (43,66). Anti-biofilm activity of phages has also been reviewed extensively in literature (67–69). Moreover, employing multiple phages as a phage cocktail is advantageous

to target multiple pathogens in polymicrobial infections and combat emerging resistance as bacteria adapt (43,70).

1.4.2 Phage therapy pipeline

Broadly, the pipeline for phage therapy in clinical practice can be summarized in four main steps: i) isolation and identification of pathogenic bacteria causing the infection; ii) isolation and selection of phages effective against the pathogenic bacteria; iii) production, storage, and transportation of phage preparations; and iv) administration to the patient (71). While phage therapy has been extensively discussed in literature previously, this review will focus on phage isolation and selection as it is the focus of this study (66,71,72).

1.4.2.1 Isolation of phages for therapeutic use

Isolation of phages that are suitable for therapeutic use is an essential step in the phage therapy pipeline. A timeline of this process – from phage discovery to identification of lifecycle by sequencing – is summarized in Figure 2. Traditional methodologies generally involve two main steps for phage discovery – enrichment and purification – with variations to techniques depending on the environmental sample, host organism, and growth conditions (73–75). In brief, enrichment involves incubating an environmental sample with the target host bacterium to amplify the number of phages present in the sample (72). This is followed by directly plating the enriched sample onto a lawn of the host bacterium to screen for any clear zones of lysis, which suggests the presence of virulent phages (75). In contrast, turbid zones of lysis have traditionally been used as an indicator – albeit, not a definitive one – for temperate phages, allowing researchers to exclude these from further characterization work, as discussed in further detail below (72). Phages are then purified from enriched samples by passaging a single plaque for a minimum of three rounds against the host bacterium, thus ensuring a pure stock of phage for

downstream propagation and concentration (75). Concentrated phage preparations are then used for characterization assays, including whole genome sequencing (WGS) to identify the presence of lysogeny markers such as integrase genes, *att* sites, repressors, and/or excisionases. Evidently, this process is significantly labour-intensive and time-consuming, with one study reporting a range of 28 to 386 days between the time request for phage therapy and administration to the patient (76). This was further apparent in another study, where only three phages were deemed suitable for therapeutic use in a cystic fibrosis patient after screening 100 environmental samples and more than 10 000 phages infecting *Mycobacterium smegmatis* (77).

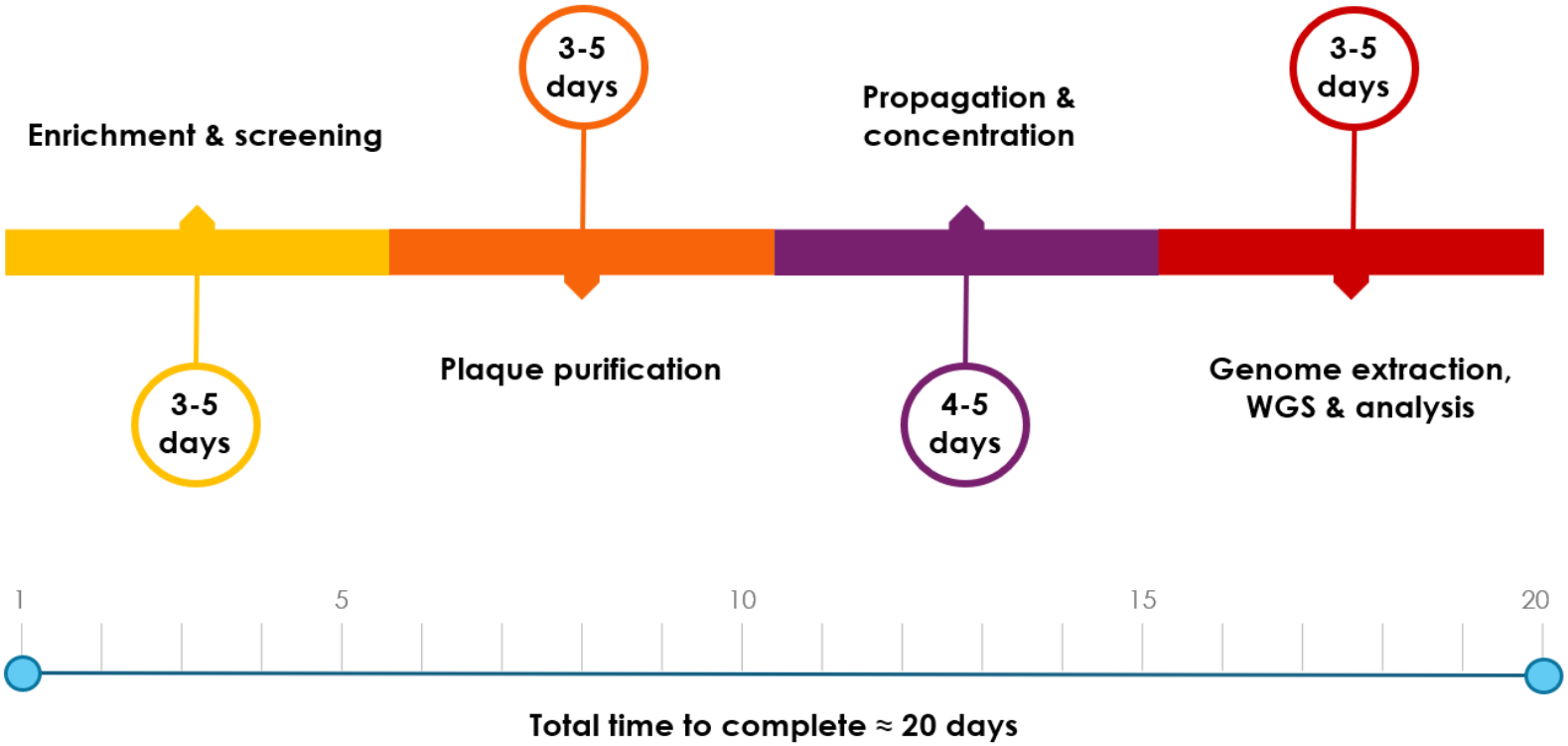

Figure 2. An estimated timeline for the isolation, purification, propagation, and sequencing of newly isolated phages from environmental samples.

1.4.2.1.1 High-throughput phage isolation

Rapid phage isolation is one of the greatest bottlenecks in the phage therapy pipeline, often increasing timescales to administration. Successful efforts have been made to streamline phage isolation by performing a high number of low-volume enrichments in multi-well microtiter plates and monitoring bacterial clearance over time in a plate assay or by screening for zones of clearing on bacterial lawns (73–75). These studies revealed significant reductions in hands-on time, total experimentation time, and monetary costs in comparison to traditional methods (74,75). However, low-volume enrichments may not be successful if the concentration of phage in the environmental samples is low; a limitation that may be ameliorated by increasing the enrichment volume and/or using concentrated samples (74,75). Efforts have also been made to simplify experimental set-up of traditional plating techniques, claiming more than 45 environmental samples can be readily screened per hour by reducing the amount of materials needed (78). However, these approaches did not account for the need to enrich an extensive collection of environmental samples (78).

1.4.2.2 Characterization and selection of ideal phages for therapeutic use

Phage characterization is an essential step in making careful and informed selection of suitable phages for therapeutic use. Some properties that are commonly assessed include host range, lifecycle, virulence, transduction potential, and genotype (72,79,80). Ideal characteristics include – but are not limited to – a broad host range to target more than one strain or species; high virulence to ensure clearance of the pathogenic bacteria; and the absence of genes encoding integrases, toxins, virulence factors, and antibiotic resistance (79,80). Phages should also be stable in storage, and easily produced in large quantities under laboratory and industrial conditions (72). Above all, the use of strictly virulent phages has been significantly underscored

to ensure lysis of target bacterial pathogens. Conversely, the use of temperate phages has been highly discouraged due to their potential to enter the lysogenic cycle, resulting in integration into the host genome and replication without lysing the target bacterial pathogen (43). Additionally, their ability to mediate transfer of harmful host-encoded genes through specialized transduction, modify host phenotypes by lysogenic conversion, and provide immunity to superinfection by the same or similar phages has made the use of temperate phages largely inadvisable (81).

1.5 Molecular detection of bacteriophages

Due to the extreme phage diversity discussed earlier in this review, no universal marker – synonymous with 16S rRNA in bacterial genomes – exists ubiquitously across all phages that would allow for rapid phage identification and selection (47). In clinical settings, this presents a challenge as early identification of phage lifecycles would expedite the phage therapy pipeline. However, earlier reports identified conserved sequences amongst closely related phages, which in turn were utilized as molecular markers, or signature genes, to explore phage diversity and relatedness (64,82–84).

1.5.1 Signature genes used for PCR-based detection

The identification of signature genes has accelerated preliminary characterization of phages by simple use of PCR. For example, detection of the major capsid protein was previously performed to identify established groups of strictly virulent phages that infect *Salmonella enterica*, *Escherichia coli*, and *Erwinia amylovora* by multiplex PCR (82). Another study also demonstrated that conserved regions within the tape measure protein in mycobacteriophages can be used as PCR targets for designation into phage clusters (83). Additionally, the integrase gene was previously used as a signature gene for PCR typing to evaluate prophage diversity as an

indicator for genomic diversity of *Salmonella enterica* host genomes, as well as to analyze the diversity of temperate enteric phages (64,85).

1.6 Rationale

The selection of suitable candidate bacteriophages for use in phage therapy is critical to treatment success. While at the genomic level, established guidelines strongly discourage phages that have undesirable genes such as toxins, integrases, and antibiotic resistance, the use of strictly virulent phages has been broadly emphasized (71).

At the time of conception of this project, the process for selecting candidate phages at Cytophage Technologies Inc. (CTI) followed the long-established, time-consuming, and costly sequence of events: enrichment, isolation, characterization, and selection (73,79). Additionally, a high-throughput method that employed rapid PCR screening of environmental isolates – immediately after enrichment – to selectively isolate virulent phages targeting clinically relevant *Staphylococcus* spp. and *Enterococcus* spp. had not yet been explored at CTI. Instead, avoidance of temperate phages during the initial isolation stages relied on the traditional absence of plaque turbidity as an initial indicator of virulence, followed by confirmation of lifecycle by whole genome sequencing (72).

As integration into the host genome is a hallmark feature associated with lysogeny, detection of integrases by PCR may allow for rapid detection, and therefore exclusion, of temperate phages during the early stages of isolation (58,72). This would save valuable time and resources by preventing unnecessary characterization of phages that are unsuitable for phage therapy. Therefore, merging a simple PCR-based assay with previous methods to streamline

phage isolation may significantly accelerate the isolation of suitable candidates for phage therapy, although further investigation is required (73).

1.7 Hypothesis

Phage integrases can be used as signature genes to screen out temperate phages infecting *Enterococcus* spp. and *Staphylococcus* spp. during high-throughput phage isolation.

1.8 Objectives

The objectives of this work were as follows:

1. Identify conserved sequences in integrase genes of temperate phages that infect *Enterococcus* spp. and *Staphylococcus* spp.
2. Develop and optimize a PCR-based method to target temperate phage integrase genes
3. Validate a high-throughput PCR-based method to detect temperate phages isolated from environmental samples

CHAPTER 2.0: MATERIALS AND METHODS

2.1 Collection of phage integrase nucleotide coding sequences from public databases

Nucleotide coding sequences of integrases belonging to temperate phages infecting *Enterococcus* spp. and *Staphylococcus* spp. were retrieved from NCBI GenBank using the following criteria: a) presence of complete and intact phage genomes; b) similar nucleotide length (i.e. complete coding sequence); c) E-value $\leq$ 0.05%, and; d) an integrase percent identity of $\approx$ 20-99%. In instances where few temperate phages could be collected from NCBI due to lack of publicly available data, PHASTER (PHAge Search Tool Enhanced Release) was used to detect complete and intact prophages in bacterial genomes available from NCBI, from which integrase nucleotide sequences were extracted. Additionally, the recombinase family (i.e. tyrosine or serine) of each integrase was identified using the NCBI Conserved Domain Search from their amino acid sequences.

2.2 Identification of conserved sequences in phage integrases

2.2.1 Phage cluster analysis

Due to the extensive global diversity of phage integrases, phage cluster analysis was performed to identify groups of similarly related phages for subsequent identification of conserved sequences. To perform phage cluster analysis, multiple sequence alignments (MSAs) were first performed using MAFFT (L-INS-i; gap open penalty = 3; off-set value = 0.5; scoring matrix = 200PAM) in Geneious Prime 2024.0.1 (https://www.geneious.com) by aligning all integrase nucleotide coding sequences that infect the same bacterial genus and belong to the same recombinase family to generate a percent identity matrix by pairwise comparison. The L-INS-i algorithm was selected as it is fast and accurate for alignments of < 200 sequences, a

moderate gap open penalty of 3 was selected to conserve homology without introducing too many gaps unless divergence is confidently observed, an off-set value of 0.5 was selected to avoid over-extending gaps and favour realistic biological events, and a default scoring matrix was maintained at 200PAM. Clusters were then identified using heatmaps generated with the heatmaply package in R at a percent identity threshold of $\gtrapprox$ 50%. Any clusters that contained $\leq$ 5 members were considered singletons and were excluded from analysis. Additional troubleshooting procedures that did not alter the outcome of this thesis, and were therefore excluded from discussion, were listed in Appendix A.1.

2.2.2 Identification of cluster-specific conserved sequences

MSAs were performed by aligning integrase nucleotide sequences belonging to the same cluster, as identified by heatmap analysis, to identify cluster-specific conserved sequences for primer design. MSAs were performed using MAFFT in Geneious Prime 2024.0.1 (see section 2.2.1 for parameters). Consensus sequences were generated at an 85% consensus threshold level.

2.3 Developing a PCR-based method to target temperate phage integrases

2.3.1 Cluster-specific primer design

Conserved nucleotides identified in section 2.2.2 were used to design cluster-specific degenerate primers for the purpose of detecting temperate phages using a PCR-based method (Tables 1 and 2). Conserved sequences that allowed for the lowest degeneracy scores possible were targets for primer design to encourage specific binding to integrase genes. Additionally, primers were designed to be 18-28 bp in length with a GC content of ~40-60%. In instances

where only short (≤ 18 bp) conserved sequences were identified and/or the GC content was < 40%, an EciI restriction site (GGCGGA) was added to the 5′ end of the primer.

2.3.2 Optimization of multiplex PCR conditions

Multiplex PCR (mPCR) conditions were optimized for detection of phage integrase genes using cluster-specific primers, as discussed above, and according to the GoTaq® Green Master Mix (Promega, Wisconsin, USA) standard reaction conditions. All PCRs were performed in a final reaction volume of 25 µL containing: 12.5 µL of 2X GoTaq® Green Master Mix, 1 µL of each forward primer (final concentration of 0.4 µM), 1 µL of each reverse primer (final concentration of 0.4 µM), 0.75 ng of gBlock template DNA, and nuclease-free water to a final volume of 25 µL. Primer sets targeting enterococcal temperate phages were combined into two mPCR reactions: multiplex reaction #1, targeting Ent-SA and Ent-TA, and; multiplex reaction #2, targeting Ent-SB, Ent-TB, and Ent-TC. Primer sets targeting staphylococcal temperate phages were also combined into two mPCR reactions: multiplex reaction #1, targeting Sta-SA and Sta-TA, and; multiplex reaction #2, targeting Sta-TB and Sta-TC. Short synthesized DNA sequences, herein referred to as gBlocks, were designed using the coding sequence of an integrase belonging to each cluster and were used as template DNA. All gBlocks were synthesized by Integrated DNA Technologies (Iowa, USA). PCR products were resolved by gel electrophoresis on a 2% (w/v) agarose gel with a 100-bp DNA ladder (New England Biolabs, Massachusetts, USA) at 100 V for 45 minutes in 1X Tris-acetate EDTA (Bio-Rad, California, USA). DNA fragments were visualized by staining the gel in a 3X GelRed™ solution (VWR, Pennsylvania, USA) for 20 minutes and illumination under UV light at 320 nm. Primers and optimized PCR conditions are detailed in Tables 1 and 2. Additional troubleshooting procedures

that did not alter the outcome of this thesis, and were therefore excluded from discussion, were listed in Appendix A.1.

Table 1. Cluster-specific degenerate primer sets for mPCR detection of *Enterococcus* spp. temperate phage integrase genes.

Multiplex Reacti	Cluster	Primer Name	Primer Sequence[a] (5′-3′)	Degeneracy score	T_a (°C)	Amplicon size (bp)
#1	Ent-SA	Ent-SA-F	CCGTTWCACAACCWTTYTTYAA	16	53	160
		Ent-SA-R	ACKYTWCCRSTWACVACTTC	192		
	Ent-TA	Ent-TA-F	GGCGGAACTGGWTTRAGA	4		205
		Ent-TA-R	GGCGGAYTTCRATYAWTTC	16		
#2	Ent-SB	Ent-SB-F	GTATTTCGTTATGTATCAGGAGGC	0	53	116
		Ent-SB-R	AGACTAGGTTGTTTGCAGTTCC	0		
	Ent-TB	Ent-TB-F	GGCGGAGTAYTTYAARACNTATC	32		204
		Ent-TB-R	GGCGGATADYKTTCHAKCCA	72		
	Ent-TC	Ent-TC-F	GATCRAWGTMTGGATACCTTACGC	8		405
		Ent-TC-R	CTCATTCCAGATTTTAGCAGAGT	0		

[a]EciI restriction sites are in red text

Table 2. Cluster-specific degenerate primer sets for mPCR detection of *Staphylococcus* spp. temperate phage integrase genes.

Multiplex Reaction	Cluster	Primer Name	Primer Sequence[a]	Degeneracy score	T_a (°C)	Amplicon size (bp)
#1	Sta-SA	Staph-SA-F	GTWACRTTAGTWGGYGCTATGGC	16	48	289
		Staph-SA-R	CTCTYGTTATWGHYCTATCTTCCCA	24		
	Sta-TA	Staph-TA-F	CDCGWWCWRTKTGGKCAGG	192		225
		Staph-TA-R	CTTCKCCDAYYCTHARACC	144		
#2	Sta-TB	Staph-TB-F	GGCGGAGGHTTYARAACDAA	36	50	352
		Staph-TB-R	GGCGGACNTCRTAYAMYGC	64		
	Sta-TC	Staph-TC-F	CTTTAACDGGWATGCGYATAGG	12		256
		Staph-TC-R	CCTCTRTCAAYATARCTYGRWTYCCA	128		

[a]EciI restriction sites are in red text

2.4 High-throughput phage isolation

2.4.1 Bacterial strains and culture media

Enterococcus faecalis 8413 was received from the Felix d'Herelle Reference Center for Bacterial Viruses, University of Laval, Canada and was used as a representative strain for isolation of enterococcal phages. *E. faecalis* 8413 was routinely cultured and maintained on Tryptic Soy broth (TSB) or Tryptic Soy agar (TSA) at 37°C for 18 hours at 200 rpm, unless otherwise stated.

Staphylococcus aureus RP29 was received from the Alberta Microbiological Repository in Calgary, Canada and was used as a representative isolation host for staphylococcal phages. *S. aureus* RP29 was routinely cultured and maintained on TSB or TSA at 37°C for 18 hours at 200 rpm, unless otherwise stated.

2.4.2 Phage enrichment and screening

To enrich a large number of environmental samples for infective phages against the desired hosts, high-throughput phage enrichment was adapted as previously described (73). A list of environmental samples screened in this study were detailed in Appendix A.2. Briefly, overnight cultures of each bacterial strain were prepared by inoculating one colony into 5 mL of TSB according to the culturing conditions described above. Then, 425 µL of 4X TSB supplemented with 8 mM $CaCl_2$, 75 µL of overnight culture, and 1 mL of environmental samples were aliquoted to the necessary number of wells of a deep-well 96-well plate. Sterile saline-magnesium (SM) buffer (100 mM NaCl, 8 mM $MgSO_4 \cdot 7H_2O$, 50 mM Tris-HCl, pH 7.5) was used in place of an environmental sample to serve as a negative control. Enrichments were then incubated at 37°C for up to 72 hours at 120 rpm. Following incubation, 200 µL of the enriched

environmental samples were transferred to a 96-well plate with 0.45 µm filters and centrifuged at 3500 x g for 5 minutes to remove debris. The filtrate was then screened for the presence of infective phages against the isolation host by performing a spot test using the double agar overlay method (86). This was carried out by adding 200 µL of overnight culture to 4 mL to the appropriate molten soft agar overlay (*E. faecalis* 8413, TSB + 0.6% (w/v) agar + 2 mM $CaCl_2$; *S. aureus* RP29, Luria-Bertani broth (LB) + 0.4% (w/v) agar + 2 mM $CaCl_2$ for *S. aureus* RP29) and pour plating over petri dishes containing solidified underlay containing TSA + 1.5% (w/v) agar, and allowing the soft agar overlay to solidify for approximately 10 minutes. 10 µL of filtered enrichment was then spotted on the overlay and allowed to dry. Once dried, plates were incubated overnight at 37°C for *E. faecalis* 8413 or 30°C for *S. aureus*. Following incubation, plates were inspected for lytic activity by the presence of clear zones, which were considered positive for infective phages.

2.4.3 Phage isolation and purification

Enriched samples that screened positive for the presence of infective phages were plated for single, isolated plaques using the double agar overlay method as previously described (87). To do this, filtered enrichments were serially diluted 10-fold in SM buffer to 10^{-6} to a final volume of 100 µL. 75 µL of dilutions 10^{-1} to 10^{-6} were then mixed with 66 µL of overnight culture of the appropriate host bacterium. 1.5 mL of the appropriate molten soft agar overlay was then added to each tube and pour-plated in each section of segmented Y-plates (VWR, Pennsylvania, USA) containing the solidified TSA agar underlay specified in section 2.4.2. The overlay was allowed to solidify followed by overnight incubation at 37°C for *E. faecalis* 8413 or 30°C for *S. aureus* RP29. A total of ten well-separated plaques were initially picked and resuspended in 50 µL of SM buffer. Resuspended plaques were then serially diluted, plated with

their respective host bacterium, and picked again using the same process as described above. After three rounds of purification, all remaining phages were deemed purified. Phages were then titrated for plaque-forming-units (PFU) per mL by serially diluting 10-fold in SM buffer and spot plating 10 µL of each dilution in triplicate on soft agar overlay containing the bacterial host, as previously described (86). PFUs were counted after plates were incubated overnight at the appropriate host temperature.

2.4.4 Bacteriophage propagation and concentration

2.4.4.1 Soft-agar overlay propagation

To generate working stocks of isolates, phage propagation in soft agar overlay was conducted as previously described (88). Purified plaque picks were diluted in SM buffer to the highest dilution factor that gave a full clearing – and not isolated plaques – during phage titration, mixed with 200 µL of an overnight culture of host bacterium in 4 mL of appropriate molten overlay, and pour plated onto a TSA underlay plate. After overnight incubation at the appropriate host temperature, 4 mL of SM buffer was distributed onto the surface of the overlay and incubated at room temperature for ~ 2 hours at 50 rpm to allow phages to diffuse out of the overlay. SM buffer and overlay were then collected into a 50 mL falcon tube, centrifuged at 4000 *x g* for 20 min at 4°C to remove debris, followed by filter sterilization of the phage-containing supernatant with a 0.22 µm filter. Phages were then enumerated for PFU/mL as described in section 2.4.3.

2.4.4.2 Liquid propagation and concentration

To generate high titer ($\geq 10^8$ PFU/mL) phage preparations for downstream genome extractions and sequencing, small-scale propagations of each phage were carried out in liquid media. To start, overnight cultures of the desired bacterial host were prepared as previously described (section 2.4.2). 2.5 mL of overnight cultures (*E. faecalis* 8413, ~5 x 10^7 CFU/mL; *S. aureus* RP29, 3.67 x 10^7 CFU/mL) were then infected 1:1 (v/v) at a multiplicity of infection (MOI) of ~10^{-6}, and incubated statically for 5 minutes at room temperature to encourage phage adsorption. The entire volume of phage-host mixture was then transferred to 100 mL TSB supplemented with 2 mM $CaCl_2$ and incubated at 37°C overnight with agitation at 120 rpm to allow for phage propagation. After incubation, bacterial debris was removed by centrifugation at 4000 x g for 20 minutes at 4°C followed by filter sterilization of the supernatant with a 0.22-µm Nalgene™ Rapid-Flow™ bottle top filter (ThermoFisher, Massachusetts, USA). The entire volume of the phage-containing filtrate was then concentrated by ultrafiltration using Pall® 100K Centrifugal Device filters (VWR, Pennsylvania, USA) at 4000 x g until the final volume was ~ 1 mL (88). The concentrated phage was then washed twice with ~15 mL of sterile SM buffer at 4000 *x g* for 20 min, or until the final volume was ~1 mL, in order to perform a buffer exchange. The concentrated phages were then enumerated for PFU/mL as described in section 2.4.3 to ensure the titers were high enough for DNA extraction ($\geq 10^{10}$ PFU/mL).

2.5 Whole genome sequencing and bioinformatic analyses

2.5.1 Extraction of genomic phage DNA

2.5.1.1 Phenol-chloroform-isoamyl alcohol method

Extraction of genomic phage DNA for whole genome sequencing was adapted from and conducted as previously described elsewhere (89). To start, 400 µL of high titer phage ($\geq 10^{10}$ PFU/mL) was pre-treated with 10 µL of DNase I (2000 U/mL) and 10 µL of RNase I_f (50 000 U/mL) at 37°C for 1 hour to remove host nucleic acids. 400 µL of 2% SDS containing 10 mg/mL of proteinase K (New England Biolabs, Massachusetts, USA) was then added and incubated for 1 hour at 56°C to lyse phages, followed by incubation on ice for 5 min. 400 µL of phenol:chloroform:isoamyl alcohol (25:24:1; Thermo Fisher, Massachusetts, USA) was then added to the phage lysates, mixed by inversion for 5 min, and centrifuged at 16 000 *x g* for 10 min to allow for phage separation. The aqueous layer was then collected and mixed twice with 1 volume chloroform (Thermo Fisher, Massachusetts, USA) for 5 min, followed by centrifugation at 16 000 *x g* for 10 min to remove residual phenol. After collecting the washed aqueous layer, DNA precipitation was performed by adding 0.1 volumes of 3 M sodium acetate (pH 5.2) and 0.7 volumes of isopropanol, mixing by inversion for ~ 5 min, and incubating overnight at -20°C. Genomic DNA was then pelleted at 17 900 *x g* for 10 min at 4°C. The resulting pellet was then washed with 500 µL of cold 70% (v/v) ethanol at 17 900 *x g* for 5 min at 4°C, air dried, and resuspended in nuclease free water. DNA concentrations were measured using the QuantiFluor® ONE dsDNA (Promega, Wisconsin, USA) and Nanodrop spectrophotometer systems (Thermo Fisher, Massachusetts, USA).

2.5.1.2 Protein precipitation method

When access to a fumehood was not available, extraction of genomic phage DNA was adapted from the Wizard™ whole blood DNA extraction kit protocol (Promega, Wisconsin, USA). Firstly, 300 µL of high titer phage ($\geq 10^{10}$ PFU/mL) was pretreated with DNase I and RNase I_f as detailed in section 2.6.1.1. Then, 300 µL of 4% SDS and 50 µL of proteinase K (20 mg/mL) (New England Biolabs, Massachusetts, USA) were added and incubated for 1 hour at 56°C to lyse phages. Phage lysates were then chilled on ice for 5 minutes, treated with 100 µL of Protein Precipitation Solution® (Promega, Wisconsin, USA), vortexed vigorously for 20 seconds, and chilled on ice for 30 minutes to promote protein precipitation. Proteins were then removed by centrifugation at 16 000 x g at 4°C for 10 minutes or until all protein had precipitated. DNA was precipitated from the supernatant with 300 µL of room temperature isopropanol for 3 minutes with mixing by inversion, followed by centrifugation at 17 900 x g for 10 minutes at 4°C to pellet the DNA. The DNA pellet was then washed with 500 µL of cold 70% (v/v) ethanol and centrifuged at 17 900 x g for 5 minutes at 4°C. The resulting DNA pellet was resuspended in nuclease-free water overnight at 4°C. DNA was quantified as described in section 2.6.1.1.

2.5.2 Whole genome sequencing and bioinformatic analysis

To determine the presence – or absence – of integrase genes and predict lifecycles of environmental phage isolates, whole genome sequencing of purified genomic phage DNA was carried out according to the Nanopore Native Barcoding Kit 24 V14 protocol (Oxford Nanopore). Sample libraries were run on the Oxford Nanopore Flongle sequencing platform for 18 hours with a minimum Q-score of 6 and minimum fragment length of 200 bp. Barcodes were trimmed using the MinKNOW software during sequencing runs. Single-molecule reads were

assembled *de novo* using the flye assembly strategy and annotated to predict open reading frames (ORFs) using the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) computational tools (https://www.bv-brc.org/). Assembled contigs were also subjected to BLASTN search using default parameter settings to predict the identity of isolated phages based on percent identity of whole genomes and major capsid protein (MCP) (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Assembled contigs were further analyzed using PhageAI (https://www.phage.ai/) to predict the phage lifecycle.

2.6 Plaque PCR screening to detect temperate phages

Using mPCR and gel electrophoresis conditions optimized in section 2.3.2, environmental phage isolates were screened for the presence of integrase genes. To extract crude phage DNA, 10 µL of plaque picks were boiled at 95°C for 10 minutes. 2.5 µL of boiled plaque picks were used as template DNA in the PCR reactions and the volume of nuclease-free water was adjusted accordingly. Any band sizes that did not correspond to the target amplicon sizes were considered negative (82).

2.6.1 Sanger sequencing of PCR products

The specificity of the primers was assessed by Sanger sequencing to confirm that the identities of PCR products were from integrases and not of other off-target sequences. Based on the number of samples, a subset of PCR products that screened positive for the presence of integrase genes were purified using the DNA Clean-Up & Concentrator-25 kit (Zymo Research, California, USA). PCR products that were too short to be directly sequenced ($\lesssim$ 300 bp) were inserted into the pCR™4-TOPO™ TA vector according to the TOPO™ TA Cloning™ kit for

Sequencing with One Shot™ TOP10 Chemically Competent *E.coli* (Invitrogen, Massachusetts, USA). Transformants were screened for plasmids containing the correct-sized insert by colony PCR using the GoTaq Green Master Mix (Promega, Wisconsin, USA) standard conditions and gel electrophoresis, as described in section 2.3.2. For colony PCR, primers M13(-41)-F (GGTTTTCCCAGTCACGAC) and M13(-48)-R (AGCGGATAACAATTTCACAC) were used at final concentrations of 0.2 μM and annealing temperature 51.3°C. Plasmid DNA was then isolated from positive transformants using the QIAprep Miniprep kit (Qiagen, Hilden, Germany). DNA was quantified using the QuantiFluor® ONE dsDNA (Promega, Wisconsin, USA) and Nanodrop spectrophotometer systems (Thermo Fisher, Massachusetts, USA). All PCR products were sequenced by Genewiz (Azenta Life Sciences, Massachusetts, USA). Sequenced PCR products were then subjected to a BLASTN search using default parameter settings to determine the sequence identity.

CHAPTER 3.0: RESULTS

3.1 Phage-cluster analysis of integrase genes

Nucleotide sequences of temperate phage integrase genes infecting *Enterococcus* spp. (n = 70) and *Staphylococcus* spp. (n = 72) were collected from NCBI GenBank (Appendix A.3 and Appendix A.4). Phage cluster analysis was conducted to find conserved sequences in temperate phage integrase sequences collected as detailed in section 2.1.1. Clusters were identified using heatmaps from percent identity matrices at a ≳ 50% identity threshold level (Figure 3). Overall, serine integrases of temperate phages that infect *Enterococcus* spp. were sorted into two clusters (Figure 3A) – denoted as Ent-SA and Ent-SB – whereas three clusters were identified from tyrosine integrases, denoted as Ent-TA, Ent-TB, and Ent-TC (Figure 3B). In contrast, one cluster of serine integrases was identified from temperate phages that infect *Staphylococcus* spp. (Figure 3C) – named Sta-SA – whereas three clusters were identified among tyrosine integrases, denoted as Sta-TA, Sta-TB, and Sta-TC (Figure 3D).

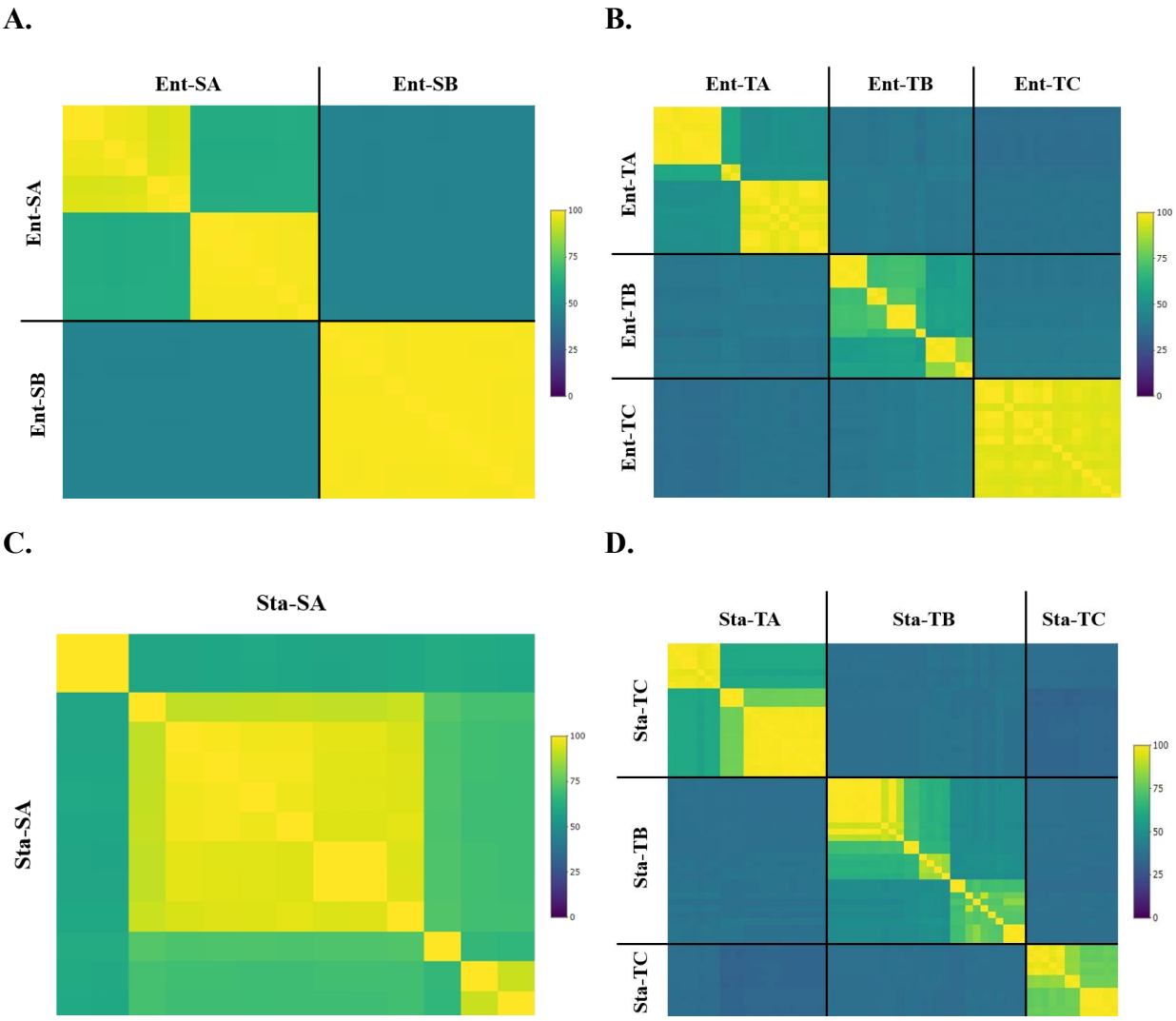

Figure 3. Phage-cluster relationships of integrases from temperate phages that infect *Enterococcus* spp. (A and B) and *Staphylococcus* spp. (C and D). Nucleotide coding sequences of serine (**A and C**) and tyrosine (**B and D**) integrases were collected from NCBI GenBank and aligned using MAFFT to calculate pairwise percent identities. Percent identity matrices were visualized with heatmaps using the heatmaply package in R. The percent identity threshold for any given cluster was $\gtrapprox$ 50%.

3.2 PCR optimization for detection of temperate phage integrase genes

Optimization of mPCR conditions was performed using gBlocks as template DNA for integrase gene sequences, and cluster-specific primers (Tables 1 and 2). For enterococcal phage integrases, sensitive and specific amplification was observed by optimized mPCR conditions using gBlocks as template DNA (Figure 4). In multiplex reaction #1, primers targeting integrase sequences belonging to Ent-SA and Ent-TA resulted in target band sizes of 160 bp and 205 bp, respectively, whereas amplification was not observed from template DNA belonging to clusters Ent-SB, Ent-TB, and Ent-TC (Figure 4A). Conversely, primers targeting gBlocks representing integrase clusters Ent-SB, Ent-TB, and Ent-TC in multiplex reaction #2 resulted in correct amplification of band sizes 116, 204, and 405 bp, respectively, whereas no amplification from Ent-SA and Ent-TA was observed (Figure 4B).

For staphylococcal phage integrases, primers targeting integrase genes belonging to Sta-SA and Sta-TA in multiplex reaction #1 resulted in amplification of target band sizes 289 bp and 225 bp, respectively, whereas no amplification was observed from template DNA belonging to clusters Sta-TB and Sta-TC (Figure 5A). In contrast, primers targeting integrase genes belonging to clusters Sta-TB and Sta-TC in multiplex reaction #2 resulted in amplification of target band sizes 352 bp and 256 bp, respectively, whereas no amplification from sequences belonging to Sta-SA and Sta-TA was observed (Figure 5B).

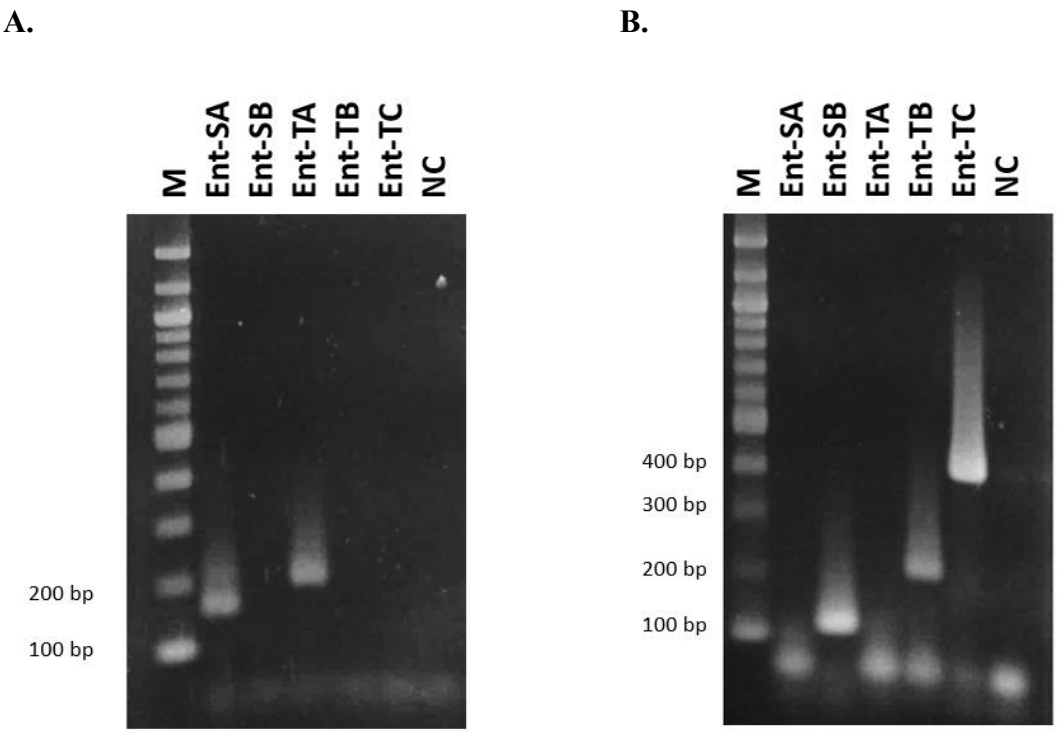

Figure 4. Agarose gel electrophoresis of mPCR products using cluster-specific primers to target integrase gene sequences of temperate phages infecting *Enterococcus* spp. Multiplex PCRs were separated into two reactions using primers specific for clusters Ent-SA and Ent-TA in multiplex reaction #1 (**A**), and Ent-SB, Ent-TB and Ent-TC in multiplex reaction #2 (**B**). PCR products were amplified using gBlocks, as template DNA, representative of temperate phage integrase genes belonging to clusters Ent-SA, Ent-TA, Ent-SB, Ent-TB, and Ent-TC. A 100-bp ladder (M) and negative control (NC) were also resolved by gel electrophoresis on a 2% agarose gel at 100 V for 45 min.

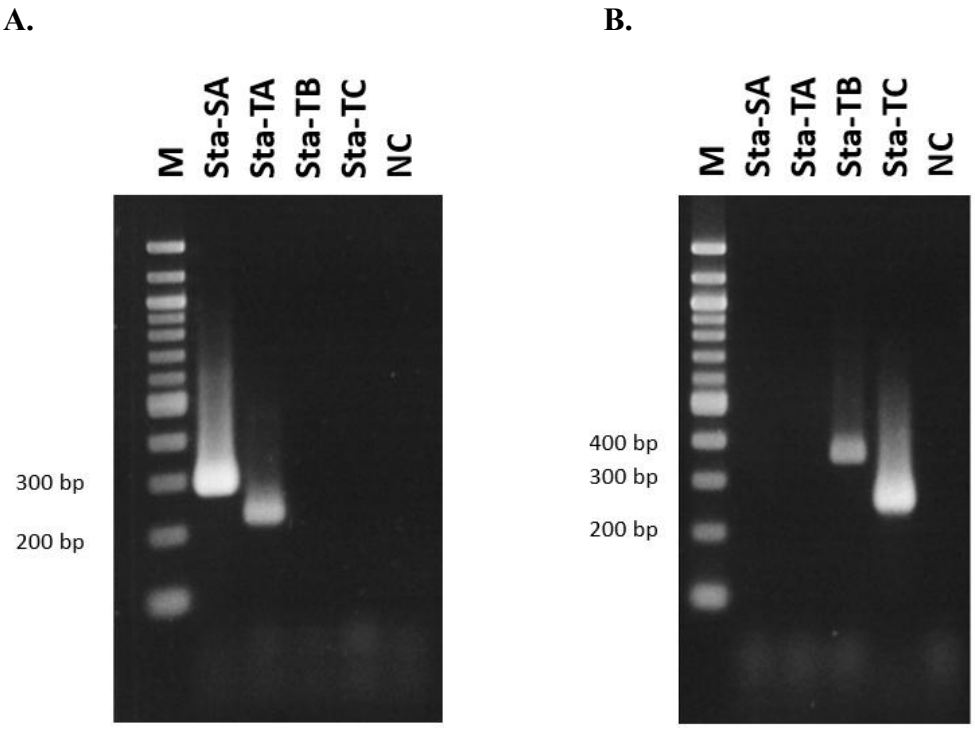

Figure 5. Agarose gel electrophoresis of multiplex PCR products using cluster-specific primers to target integrase gene sequences of temperate phages infecting *Staphylococcus* spp. Multiplex PCRs were separated into two reactions using primers specific for clusters Sta-SA and Sta-TA in multiplex reaction #1 (**A**), and clusters Sta-TB and Sta-TC in multiplex reaction #2 (**B**). PCR products were amplified using gBlocks, as template DNA, representative of temperate phage integrase genes belonging to clusters Sta-SA, Sta-TA, Sta-TB, and Sta-TC. A 100-bp ladder (M) and negative control (NC) were also resolved by gel electrophoresis on a 2% agarose gel at 100 V for 45 min.

3.3 High-throughput phage isolation

3.3.1 Isolation of phages infecting *E. faecalis* 8413

To isolate phages targeting host *E. faecalis* 8413, ten environmental samples were enriched. However, only five environmental samples (CIE9, CIE23, CIE49, CIE51, and CIE102) screened positive for the presence of infective phages based on spot testing (data not shown). From the ten plaques picked from enrichments, seven plaques were purified and propagated on a larger scale for subsequent genome extraction and whole genome sequencing. The list of isolated of phages and their environmental sources are detailed in Table 3.

3.3.2. Isolation of phages infecting *S. aureus* RP29

Infective phages against *S. aureus* RP29 were detected from twelve enriched environmental samples (data not shown). From the ten plaques picked from enrichments, nine were purified and propagated on a larger scale for subsequent genome extraction and whole genome sequencing. The list of isolated phages and their environmental sources are detailed in Table 3.

Table 3. Environmental phage isolates using hosts *E. faecalis* 8413 and *S. aureus* RP29.

Host bacterium	Phage ID	Environmental sample	
		Sample #	**Sample source**
E. faecalis 8413	E9-1	CIE9	Sewer #3
	E23-1	CIE23	Wastewater plant #1
	E23-2	CIE23	Wastewater plant #1
	E49-2	CIE49	Hospital wastewater #2
	E51-2	CIE51	Hospital wastewater #4
	E102-1	CIE102	Sewage #2
	E102-2	CIE102	Sewage #2
S. aureus RP29	SA2-1	CIE2	Sewer #1
	SA28-2	CIE28	Wastewater plant #3
	SA48-2	CIE48	Hospital wastewater #1
	SA50-2	CIE50	Hospital wastewater #3
	SA52-2	CIE52	Medical center wastewater #1
	CIP93	CIE7	River water
	CIP94	CIE9	Sewer #3
	CIP95	CIE24	Wastewater plant #2
	CIP98	CIE23	Wastewater plant #1

3.4 Whole genome sequencing and bioinformatic analysis

Whole genome sequencing and bioinformatic analysis were carried out to identify the presence of integrase genes and predict the lifecycle for each isolated enterococcal phage. Sequencing results were summarized in Table 4.

Annotation results of enterococcal phage isolates did not detect homology between any open reading frames and integrases or site-specific recombinases. Additionally, PhageAI predicted that the lifecycle of each phage was virulent with $\geq$ 96% confidence. Conversely, annotation results of each staphylococcal phage detected an integrase in all isolates, except for CIP95. Analysis using PhageAI suggested that the lifecycles of all isolates were temperate, with confidence ranging from ~74-99%, except for CIP95 which was predicted to be virulent (~91% confidence). The identities of each phage based on whole genome and MCP sequences were also determined and are detailed in Appendix A.5.

Table 4. Presence of integrase genes and predicted lifecycles of sequenced phage isolate genomes infecting *E. faecalis* 8314 and *S. aureus* RP29.

Host bacterium	Phage ID	Integrase gene present[a]	Predicted lifecycle[b]	
			Lifecycle	Confidence (%)
E. faecalis 8413	E9-1	No	Virulent	99.48
	E23-1	No	Virulent	97.70
	E23-2	No	Virulent	96.60
	E49-2	No	Virulent	99.04
	E51-2	No	Virulent	96.31
	E102-1	No	Virulent	97.87
	E102-2	No	Virulent	97.62
S. aureus RP29	SA2-1	Yes	Temperate	95.55
	SA28-2	Yes	Temperate	98.07
	SA48-2	Yes	Temperate	83.24
	SA50-2	Yes	Temperate	74.45
	SA52-2	Yes	Temperate	98.90
	CIP93	Yes	Temperate	88.05
	CIP94	Yes	Temperate	97.25
	CIP95	No	Virulent	91.08
	CIP98	Yes	Temperate	90.93

[a]Annotated using BV-BRC; [b]Predicted using PhageAI

3.5 PCR screening of environmental phage isolates

3.5.1 Plaque PCRs of environmental phage isolates infecting *E. faecalis* 8413

PCR screening of plaque picks of enterococcal phage isolates was carried out using optimized conditions detailed in section 2.3.2. PCR screening using multiplex reaction #1, containing primers specific for Ent-SA and Ent-TA, did not result in amplification from the plaque picks (lanes 1-7) or the bacterial control (lane 8) (Figure 6A).

In contrast, PCR screening using multiplex reaction #2 (Figure 6B), containing primers specific for Ent-SB, Ent-TB, and Ent-TC, resulted in amplification of a ~405 bp product in all plaque picks (lanes 1-7) and the bacterial control (lane 8). Positive controls for Ent-SB, Ent-TB, and Ent-TC resulted in amplification of bands that were 116 bp, 204 bp, and 405 bp (lanes 9-11), respectively, whereas no amplification was observed in the negative controls of SM and nuclease-free water (lanes 12 and 13).

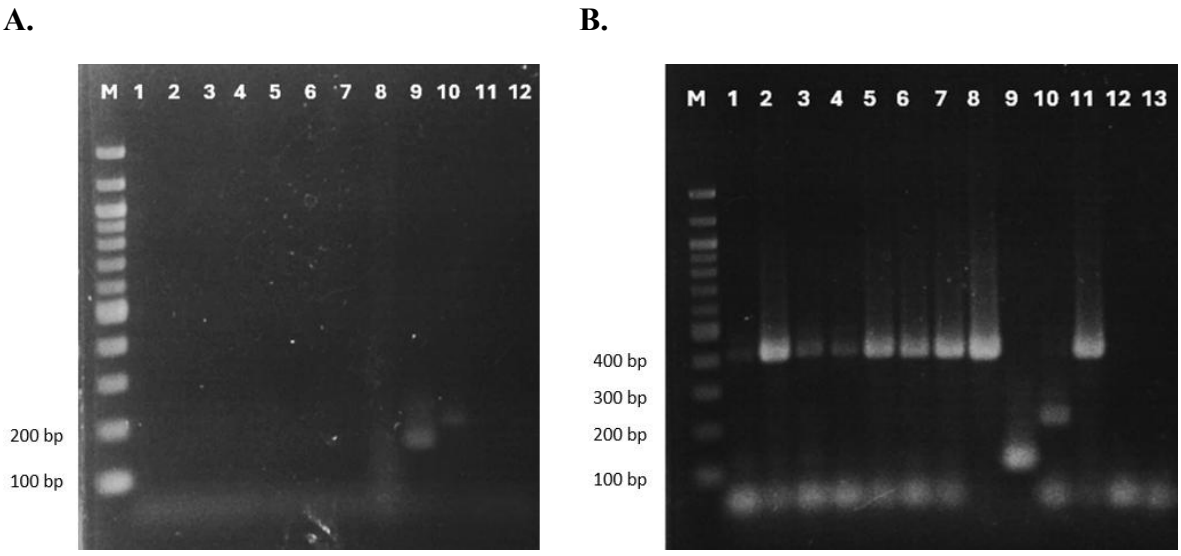

Figure 6. mPCR screening of environmental phage isolates infecting *E. faecalis* 8413. Optimized mPCR reactions #1 (**A**) and #2 (**B**) were used to screen plaque picks of environmental phage isolates E9-1, E23-1, E23-2, E49-2, E51-2, E102-1, and E102-2 (lanes 1-7). A bacterial control of the isolation host, *E. faecalis* 8413, was also screened for presence of prophages (lane 8). Positive controls of gBlocks belonging to clusters Ent-SA and Ent-TA (**A**, lanes 9-10), or Ent-SB, Ent-TB, and Ent-TC (**B**, lanes 9-11) were also amplified by mPCR. A 100-bp ladder (lane M), and negative controls using SM buffer (**A**, lane 11; **B**, lane 12) and nuclease-free water (**A**, lane 12; **B**, lane 13) were also included. PCR products were resolved by gel electrophoresis in a 2% agarose gel at 100 V for 45 min.

3.5.2. Plaque PCRs of environmental phage isolates infecting *S. aureus* RP29

PCR screening of plaque picks of staphylococcal phage isolates was carried out according to optimized conditions detailed in section 2.3.2 (Figure 7). In all staphylococcal phage isolates, plaque PCR screening using multiplex reaction #1, containing primers specific for Sta-SA and Sta-TA, resulted in amplification of two bands of ~225 bp and ~289 bp (Figure 7A; lanes 1-9). Two bands of ~250 bp and ~289 bp were also observed in the bacterial control (lane 10). Positive controls for Sta-SA and Sta-TA resulted in amplification of bands that were 225 bp and 289 bp, respectively (lanes 11 and 12). No amplification was observed in the SM and nuclease free water negative controls (lanes 13 and 14).

PCR screening of plaque picks using multiplex reaction #2 resulted in amplification from phages SA2-1, SA52-2, CIP94, CIP95, and CIP98 of various band sizes (Figure 7B; lanes 1, 5, 7-9). Various band sizes were also observed in the bacterial control (lane 10). However, none of these band sizes corresponded with those observed in the Sta-TB (352 bp) and Sta-TC (256 bp) positive controls (lanes 11 and 12). No amplification was observed in the SM and nuclease free water negative controls (lanes 13 and 14).

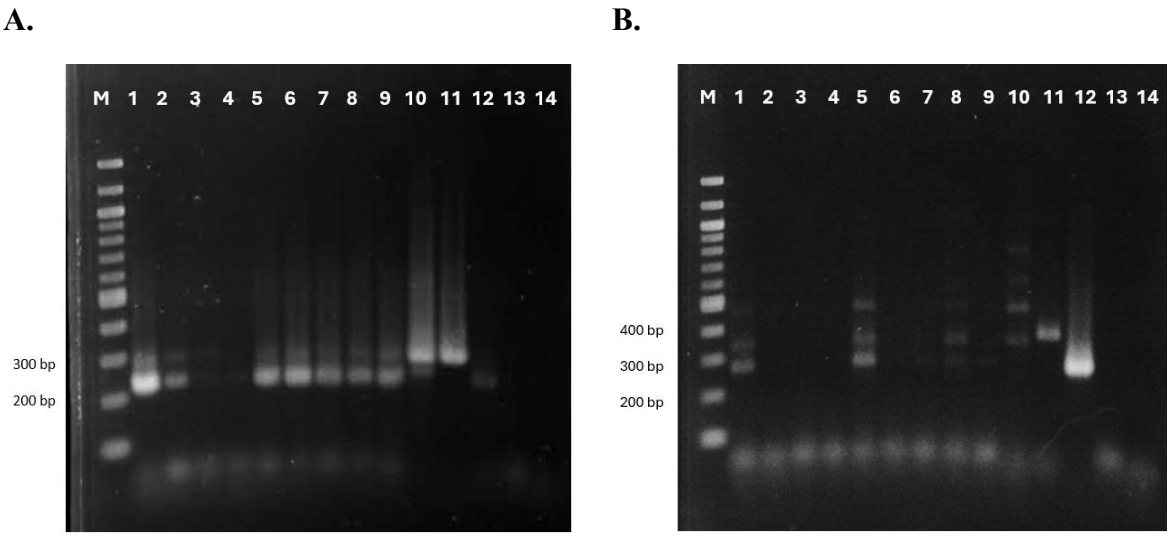

Figure 7. mPCR screening of environmental phage isolates infecting *S. aureus* RP29. Optimized mPCR reactions #1 (**A**) and #2 (**B**) were used to screen plaque picks of environmental phage isolates SA2-1, SA28-2, SA48-2, SA50-2, SA52-2, CIP93, CIP94, CIP95, and CIP98 (lanes 1-8). A bacterial control of the isolation host, *S. aureus* RP29, was also screened for presence of prophages (lane 9). Positive controls of gBlocks belonging to clusters Sta-SA and Sta-TA (**A**, lanes 10-11), or Sta-TB and Sta-TC (**B**, lanes 10-11) were also amplified by mPCR. A 100-bp ladder (lane M), and negative controls using SM buffer (**A** and **B**, lane 13) and nuclease-free water (**A** and **B**, lane 14) were also included. PCR products were resolved by gel electrophoresis in a 2% agarose gel at 100 V for 45 min.

3.6 Verifying the identity of PCR products

To confirm the specificity of PCR primers for integrase genes – and not unrelated sequences – mPCR products amplified from screening of environmental phage isolates were sequenced with Sanger sequencing and subjected to BLASTN search to determine their identity in the GenBank database.

3.6.1 PCR products from *E. faecalis* 8413 phage isolates

Based on the number of samples, a subset of PCR products from phage isolates E9-1, E23-1, E51-2, and E102-1, in addition to the bacterial control (*E. faecalis* 8413), were sequenced (Table 5). Each ~405 bp amplicon from E9-1, E23-1, E51-2, E102-1, and *E. faecalis* 8413 shared 99.04%, 98.56%, 99.04%, 98.56%, and 98.56% identity with a site-specific integrase in *Enterococcus faecalis* strain BE67 (Accession No. UYY03508.1).

PCR products were mapped to the phage and host genomes in an attempt to determine the source of amplification. All PCR products mapped to a site-specific recombinase within a prophage region within the isolation host *E. faecalis* 8413.

Table 5. Identity of PCR products amplified from a subset of *E. faecalis* 8413 environmental phage isolates.

Organism	Identity	Source	Percent identity (%)	Accession No.
E9-1	Site-specific integrase	*Enterococcus faecalis* strain BE67	99.04	UYY03508.1
E23-1	Site-specific integrase	*Enterococcus faecalis* strain BE67	98.56	UYY03508.1
E51-2	Site-specific integrase	*Enterococcus faecalis* strain BE67	99.04	UYY03508.1
E102-1	Site-specific integrase	*Enterococcus faecalis* strain BE67	98.56	UYY03508.1
E. faecalis 8413	Site-specific integrase	*Enterococcus faecalis* strain BE67	98.56	UYY03508.1

3.6.2 PCR products from *S. aureus* RP29 phage isolates

Since two PCR products (~225 bp and ~289 bp) were amplified from each environmental phage isolate, a subset of PCR products from phages SA2-1, SA28-2, and CIP95 and the bacterial host *S. aureus* RP29 were sequenced. The identities of each amplicon are summarized in Table 6.

Firstly, BLAST results identified that the ~225 bp PCR product amplified from SA2-1, SA28-2, and CIP95 all shared 95.96, 97.31, and 96.86% identity, respectively, with an integrase from *Staphylococcus* phage BUCT556 (Accession No. QSM00652.1). Additionally, the ~289 bp PCR product amplified from host *S. aureus* RP29 shared 98.28% identity with a putative site-specific recombinase in an uncultured *Caudovirales* phage (ASN70113.1). In contrast, the ~289 bp PCR product from SA2-1 and CIP95 shared 99.61% and 99.22% identity, respectively, and 30% coverage with an NAD(P)-dependent oxidoreductase (Accession No. WFN48923.1), whereas the ~289 bp PCR product amplified from SA28-2 shared 99.31% identity with an integrase from *Staphylococcus* phage IME1348_01 (Accession No. YP_010082984.1).

PCR products were also mapped to the genomes of phage isolates and *S. aureus* RP29 in an attempt to identify the source of amplification. The ~225 bp amplicon from isolates SA2-1 and SA28-2 mapped to a phage integrase within their respective genomes, whereas the ~225 bp amplicon from CIP95 and *S. aureus* RP29 did not map to the phage or host. Furthermore, the ~289 bp amplicon from SA2-1 and CIP95 mapped to a region annotated as a “hypothetical protein similar to glucose epimerase” within the host genome, whereas the ~289 bp amplicon from SA28-2 did not map to the phage or the host.

Table 6. Identity of PCR products amplified from a subset of *S. aureus* RP29 environmental phage isolates.

Organism	Amplicon size (bp)	Identity	Source	Percent identity (%)	Accession No.
SA2-1	289	NAD(P)-dependent oxidoreductase	*Staphylococcus aureus* strain 13 chromosome	99.61	WFN48923.1
	225	Integrase	*Staphylococcus* phage BUCT556	95.96	QSM00652.1
SA28-2	289	Integrase	*Staphylococcus* phage IME1348_01	99.31	YP_010082984.1
	225	Integrase	*Staphylococcus* phage BUCT556	97.31	QSM00652.1
CIP95	289	NAD(P)-dependent oxidoreductase	*Staphylococcus aureus* strain 13 chromosome	99.22	WFN48923.1
	225	Integrase	*Staphylococcus* phage BUCT556	96.86	QSM00652.1
S. aureus RP29	289	Putative site-specific recombinase	Uncultured *Caudovirales* phage	98.28	ASN70113.1

CHAPTER 4.0: DISCUSSION

While the revival of phage therapy holds promise as an alternative treatment strategy to combat antibiotic resistant pathogens, the process of isolating and selecting suitable phage candidates is a major bottleneck in the phage therapy pipeline (73). Furthermore, the demand for phages that follow a strictly virulent lifecycle has been significantly underscored by researchers, clinicians, and regulatory bodies alike (66,90,91). In an effort to streamline the process of selecting strictly virulent phages, this proof-of-concept study aimed to implement a rapid PCR-based screening assay during the early stages of high-throughput phage isolation. More specifically, this assay aimed to detect the presence of integrases from temperate phages to facilitate preliminary assignment of phage lifecycles, thus excluding them from further characterization assays in future selection of ideal phage therapy candidates. While this study revealed that specific detection of integrases by PCR may be possible when using short purified DNA sequences (Figures 4 and 5), the presence of contaminating factors in crude phage lysates is a major limitation (Figures 6 and 7).

As mentioned previously, the absence of universal markers and conserved sequences across all phages are a major hindrance for developing PCR screening assays (47). Indeed, common methods that employ global alignment of sequences to identify genetic similarities are ineffective using whole genome sequences of phages due to variations in genome length, gene content, and gene synteny (83). However, conservation identified by phage cluster analysis of single genes had previously been shown to circumvent this issue (64,82,83). Phage cluster analysis was therefore adopted in this study using integrases as molecular markers since they are a common genetic element among temperate phages (Figure 3). At a $\gtrapprox$ 50% percent identity threshold level, phage clustering analysis identified two clusters belonging to the serine

recombinase family (Ent-SA and Ent-SB) and three clusters belonging to the tyrosine recombinase family (Ent-TA, Ent-TB, and Ent-TC) among enterococcal temperate phages (Figure 3A and 3B). Additionally, one cluster belonging to the serine recombinase family (Sta-SA) and three clusters belonging to the tyrosine recombinase family (Sta-TA, Sta-TB, and Sta-TC) were identified from staphylococcal temperate phages (Figure 3C and 3D). Similar to observations in earlier studies (83,92), little homology was observed between different clusters; however, conserved sequences within each cluster were able to be identified and used as targets for cluster-specific primer design (Table 1).

Due to the lack of *Enterococcus* spp. and *Staphylococcus* spp. phages in the CTI repository, integrase sequences belonging to each cluster were synthesized as gBlocks and served as template DNA for mPCR optimization. Overall, cluster-specific primers resulted in specific and sensitive detection of target integrase sequences. This was observed in all multiplex reactions that targeted integrases from temperate phages infecting either *Enterococcus* spp. (Figure 4) or *Staphylococcus* (Figure 5), whereby each primer pair specifically amplified the PCR product from its corresponding cluster. Additionally, primer sets that were tested using integrase sequences belonging to non-target clusters did not result in amplification, and verified primer specificity. Thus, the mPCR data confirmed that cluster-specific primers can target conserved sequences within integrases and distinguish between clusters.

Previous studies demonstrated the utility of 96-well microtiter plate assays to maximize the number of low-volume enrichments performed at once in order to streamline phage isolation (73). However, combining these microtiter plate assays with subsequent PCR screening to exclude the isolation of temperate phages that infect *S. aureus* and *E. faecalis* had not yet been explored at CTI. To do this, mPCR conditions that were optimized using gBlocks were applied to

unpurified genomic DNA from boiled plaque picks of environmental isolates. Firstly, mPCR screening of phage isolates targeting host *E. faecalis* 8413 indicated that each phage encoded an integrase belonging to cluster Ent-TC (Figure 6B). However, the same band size was also observed in the bacterial control, suggesting the presence of a prophage integrated within the host genome. This indicated that amplification observed from the plaque picks had been the result of contaminating host nucleic acids from lysed bacterial cells, and not from an integrase encoded within the phage genome. This was confirmed by the absence of integrase genes within the phage genome, according to whole genome sequencing, and a predicted virulent lifecycle according to PhageAI (Table 3). Furthermore, Sanger sequencing was conducted in this study to confirm that the PCR products resulted from amplification of an integrase gene and not an off-target sequence. Based on the Sanger sequencing results, the identity of each PCR product was confirmed to be an integrase, further verifying the specificity of the PCR primers. Additionally, it was found that every amplicon was ~ 99% identical to an integrase within a prophage region in *E. faecalis* 8413. This revealed a major limitation of this assay, in which the presence of contaminating host nucleic acids from non-purified phage DNA can result in false positive results.

Similarly, screening of phage isolates infecting host *S. aureus* RP29 initially suggested that each phage isolate encoded two integrase regions belonging to clusters Sta-SA and Sta-TA using mPCR reaction #1 (Figure 7A). A band corresponding to Sta-SA was also observed in the bacterial control. This, again, suggested the presence of a prophage within the host genome and that amplification of this band from plaque picks resulted from host nucleic acid contamination. With the exception of CIP95, whole genome sequencing and analysis with PhageAI corroborated the presence of an integrase gene only belonging to cluster Sta-TA in each phage, as well as a

predicted temperate lifecycle. Conversely, no integrase genes were encoded in CIP95 and its predicted lifecycle was virulent. Therefore, the amplicons corresponding to Sta-SA from each phage were all false positives, and the presence of integrases belonging to cluster Sta-TA in all isolates – with the exception of CIP95 – were accurately detected.

PCR screening using mPCR reaction #2 resulted in amplification of various off-target band sizes from phage isolates SA2-1, SA52-2, CIP94, CIP95, and CIP98. However, none of the PCR products corresponded to either of the target clusters (Sta-TB and Sta-TC) and were deemed negative results. Off-target amplification was also observed in a similar assay using plaque picks to screen for the MCP of virulent phages, and reflects a loss of resolution of cluster-specific primers when unpurified DNA is being used in PCR (82). This may have a resulted from insufficient sequence similarity for primer binding and/or limitations using gBlocks for PCR optimization, as discussed more thoroughly below.

Since the results from mPCR reaction #1 for each *Staphylococcus* phage isolate were the same, only positive PCR products from a subset of phage isolates (SA2-1, SA28-2, and CIP95) were sequenced with Sanger sequencing in addition to the bacterial control to reduce the workload. The BLASTN results revealed that the PCR products corresponding to Sta-SA (~289 bp) from isolates SA2-1 and CIP95 shared ~96% identity with an NAD(P)-dependent oxidoreductase, while mapping to a region annotated as a "hypothetical protein similar to glucose epimerase" within the draft genome of *S. aureus* RP29. Interestingly, the same-sized PCR product from isolate SA28-2 did not map to the phage or host genomes, despite sharing ~97% identity with an integrase in *Staphylococcus* phage IME1348_01 according to BLASTN. Moreover, BLASTN results indicated that the PCR product corresponding to Sta-SA in the bacterial genome shared ~98% identity with a putative site-specific recombinase from an

uncultured *Caudovirales* phage, although could not be mapped to the host genome. Therefore, the source of amplification of this PCR product from SA28-2 and *S. aureus* RP29 is unknown and requires further investigation, although may possibly be from the environment. Nonetheless, these results demonstrated that the PCR products corresponding to cluster Sta-SA from the phage isolates were false positives and resulted from non-specific amplification. Conversely, BLASTN results indicated that the PCR product corresponding toeci Sta-TA detected from SA2-1, SA28-2, and CIP95 all shared $\geq$ 99% identity with an integrase from *Staphylococcus* phage BUCT556. The sequence of this PCR product was mapped to an integrase gene in phage genomes of SA2-1 and SA28-2, which therefore corroborated whole genome sequencing results. However, this PCR product could not be mapped to CIP95 or the host genome, and was therefore considered a false positive. Contamination was ruled out as a source of amplification given the absence of bands in the negative controls containing SM buffer and nuclease-free water. Therefore, the source of this amplification could not be determined and requires further investigation in future work.

Evidently, the presence of contaminating nucleic acids from host-encoded prophages is a major limitation of this assay. Interestingly, previous studies have exploited the presence of host-encoded prophages as means to categorize and characterize diversity of *S. aureus* strains (93–95). However, the use of purified phage DNA devoid of host nucleic acids has been widespread throughout literature in order to avoid non-specific detection during PCR or sequencing (96–99). Efforts were made to remove host nucleic acid contamination in this study by pre-treatment of plaque picks with nucleases, including DNase I, DpnI, and T5 exonuclease. Additionally, chloroform-mediated lysis of any intact resistant bacterial cells prior to nuclease treatment of the plaque picks was also explored. Yet, all attempts were unsuccessful at eliminating amplification from host contamination (data not shown). Re-optimization of PCR conditions – including

modifying primer concentrations, reducing the number of cycles, and increasing annealing temperatures – did not resolve this issue either (data not shown). Details regarding troubleshooting the removal of host nucleic acid contamination were further summarized in Appendix A.1. If this assay were to be adopted, one may cautiously assume that any same-sized bands seen in both the plaque picks and bacterial control are the result of host contamination; therefore, assignment to the corresponding cluster would be a negative result. However, one must carefully interpret these results and proceed with caution. Additionally, future work may include designing primer sets that bind to regions in the host chromosome and prophage in order to confirm the presence of host-encoded prophages as a source of contamination; although, this requires further investigation.

Interestingly, a previous study employed a similar mPCR plaque assay to target MCPs of virulent phages. Using cluster-specific primers to screen a panel of lytic and temperate phages, no false negatives or false positives were detected, indicating 100% specificity (82). Since it has been noted in literature that no cluster contains both strictly lytic and temperate phages, targeting MCPs of clusters containing only strictly virulent phages avoids the issue of amplification from host encoded-prophages (92). Similarly, it has previously been established that all known virulent *Staphylococcus* phages are classified into either the *Rountreeviridae* or *Herelleviridae* families, whereas temperate phages are not found in either (100). This was exploited in a similar PCR assay that was developed for rapid taxonomic differentiation of virulent *S. aureus* and *Klebsiella pneumoniae* phages (84). Specifically regarding staphylococcal phages, it was shown that the resolve gene was a reliable target for phages belonging to the *Herelleviridae* family, whereas the MCP was a reliable target for the *Rountreeviridae* family. Notably, no off-target amplification was observed when crude phage lysates (i.e. unpurified phage DNA) were used as

template DNA for PCR screening. This offers a promising alternative strategy that may also be applied to specifically detect strictly virulent *Enterococcus* phages, therefore avoiding the issue of amplification from host-encoded prophages when targeting integrase genes.

While this study demonstrated that gBlocks are a feasible alternative as template DNA when phage genomes are not available for initial mPCR optimization, these short purified DNA sequences did not take into account any non-specific amplification that may result from screening a whole genome. However, this was considered when interpreting PCR results by rating any off-target amplification that did not correspond with the target band sizes as a negative result for the presence of an integrase.

Lastly, as more phage genomes are sequenced, classified, and deposited into databases, it is possible that more phage clusters will be revealed over time. This challenges the accuracy of detection, whereby newly identified clusters would not be detected by the original primer sets. Therefore, phage cluster analysis may need to be revised periodically.

CHAPTER 5.0: CONCLUSION

Overall, this study demonstrated that while it is possible to detect phage integrase sequences by mPCR, this method does not accurately translate when screening crude environmental phage isolates. Phage cluster analysis of integrase genes allowed for identification of intra-cluster conservation for primer design. Cluster-specific primers that specifically targeted these conserved sequences were successfully amplified from purified short DNA sequences (i.e. gBlocks) that represented integrase genes belonging to each cluster in lieu of genomic phage DNA. When this method was applied to screening plaque picks of environmental phage isolates, integrase genes from phage isolates as well as host-encoded prophages and genes were amplified from contaminating bacterial nucleic acids. This resulted in many false positive results, challenging the specificity and accuracy of this assay. This assay may be adopted by cautiously assuming that any PCR products observed in both the plaque picks and bacterial control are the result of host contamination, and is therefore a negative result. However, future work may involve adding additional primer sets to confirm that host-encoded prophages are the source of contamination to improve this assay. Additionally, a future approach may be improved by including PCR screening of homologues found only in strictly virulent phage, although this requires further investigation. In summary, while the strategy used in the presented work did not result in the expected outcome, the development of a rapid PCR screening assay to expedite the selection of suitable phages for phage therapy is still feasible.

Appendix

Table A.1. Summary of additional troubleshooting procedures that did not alter the outcome of the thesis and therefore were excluded from discussion.

Objective	Troubleshooting procedure	Outcome
1	Cluster identification using dotplot analysis with Gepard program	Boundaries of clusters could not be well-defined. Used heatmap analysis instead
	Amino acid sequences of phage integrases were first collected (prior to using nucleotide sequences) for MSAs. Consensus sequences were back-translated to identify conserved nucleotide sequences for primer design in Geneious	Very little conservation was observed at the nucleotide level after back-translating. Any observed conservation resulted in degenerate primers with very high degeneracy scores that were not conducive to specific binding
	MSAs of amino acid sequences were attempted using PSI-Coffee and T-Coffee programs. These were selected due to their high accuracy	Run times for alignments were very long (up to ~13 hours) and not conducive to development of a high-throughput method
2	Optimized the input quantity of gBlocks to ensure that PCR products were being visualized by gel electrophoresis, and not the template DNA. This was done by resolving a range of quantities of gBlock (1, 2, 4, 8, and 16 ng) on a 2% agarose gel at 100 V for ~30 min.	At > 1 ng, the template DNA/gBlock was still visible on the gel. Therefore, it was decided to use < 1 ng (i.e. 0.75 ng) of gBlock for PCR optimization
	Annealing temperatures of a mPCR reaction were calculated by subtracting 5°C from the lowest melting temperature of a primer within a primer mix	Successful amplification
	If off-target amplification was observed in a multiplex PCR, optimization was attempted by altering primer concentrations, reducing the number of cycles, and/or performing gradient PCRs to determine the ideal annealing temperature. If off-target amplification was still present, then tried different combinations of primers sets.	Optimized PCR conditions
3	In attempt to remove host nucleic acid contamination, plaque picks were treated with DNase I alone (0.5, 2, 4, 6, 12 U) and in combination with DpnI (20 U)and T5 exonuclease (10 U). Also tried a pre-treatment with 4% (v/v) chloroform prior to nuclease treatment to lyse any intact resistant cells. Attempts were also made to reoptimize the annealing temperature, modify primer concentrations, and adjust the number of cycles	Unsuccessful at removing host contamination

Table A.2. Environmental samples used for phage isolation.

Host bacterium	Sample ID	Source
E. faecalis 8413	CIE5	Sewer #2
	CIE7	River water
	CIE9	Sewer #3
	CIE23	Wastewater plant #1
	CIE33	Sewage #1
	CIE49	Hospital wastewater #2
	CIE50	Hospital wastewater #3
	CIE51	Hospital wastewater #4
	CIE52	Hospital wastewater #5
	CIE102	Victoria, BC; Sewage
S. aureus RP29	CIE2	Sewer #1
	CIE5	Sewer #2
	CIE7	River water
	CIE23	Wastewater plant #1
	CIE24	Wastewater plant #2
	CIE28	Wastewater plant #3
	CIE33	Sewage #1
	CIE48	Hospital wastewater #1
	CIE50	Hospital wastewater #3
	CIE51	Hospital wastewater #4
	CIE52	Hospital wastewater #5
	CIE102	Sewage #2

Table A.3. Integrases from temperate phages that infect *Enterococcus* spp. collected from NCBI GenBank and their corresponding cluster based on heatmap analysis.

Recombinase family	Protein Accession No.	Organism	Cluster
Serine	ACZ63700.1	*Enterococcus* phage phiFL1A	Ent-SA
	NSN47663.1	*Enterococcus faecalis* WW_0040D	
	EGO8242093.1	*Enterococcus faecalis* FSIS11919024	
	NSV50966.1	*Enterococcus faecalis* W19-2	
	WHK13818.1	*Enterococcus faecalis* EfsC129	
	EGO8215133.1	*Enterococcus faecalis* FSIS11918124	
	PNK96056.1	*Enterococcus faecalis* FDAARGOS_347	
	BDQ48249.1	*Enterococcus faecalis* SVR2281	
	HAP2889906.1	*Enterococcus faecalis* EF13PII	
	EGO8712153.1	*Enterococcus faecalis* FSIS11814914	
	EGO5188110.1	*Enterococcus faecalis* FSIS12031524	
	EIX6402451.1	*Enterococcus faecalis* FSIS12209194	
	YP_009103095	*Enterococcus* phage EFC-1	Ent-SB
	EGO7728092.1	*Enterococcus faecalis* FSIS11922552	
	EHL2476667.1	*Enterococcus faecalis* FSIS12138974	
	EIB6793035.1	*Enterococcus faecalis*	
	BDH66130.1	*Enterococcus* sp. PLM3	
	HBI1995064.1	*Enterococcus faecalis* 2011-70-220-1	
	NSN47722.1	*Enterococcus faecalis* WW_0040D	
	HAP5605210.1	*Enterococcus faecalis* BSAC_ec488	
	HBI1792658.1	*Enterococcus faecalis* BCW_3687	
	HCQ8724748.1	*Enterococcus faecalis*	
Tyrosine	YP_003358821.1	*Enterococcus* phage phiEf11	Ent-TA
	ALO80994.1	*Enterococcus* phage vB_EfaS_IME197	
	ETJ09375.1	Prophage LambdaSa03	
	OJG48208.1	*Enterococcus gallinarum* DSM 24841	
	MCI5684531.1	*Enterococcus gallinarum* SUG181	
	NSU82283.1	*Enterococcus faecalis* FC_0034E	
	EHD3762925.1	*Enterococcus faecalis* FSIS12137147	
	HAP2895835.1	*Enterococcus faecalis* EF14PII	
	EHZ9202160.1	*Enterococcus faecalis* FSIS12140959	

	BDP47584.1	Prophage PS2	
	HAQ7626890.1	*Enterococcus faecium* BSAC_ec555	
	HAP7949763.1	*Enterococcus faecium* UW13799	
	HBL2778363.1	*Enterococcus faecium* PR28309_1	
	HBK5783432.1	*Enterococcus faecium* P-18741	
	EGP5102032.1	*Enterococcus faecium* FSIS11917240	
	EGP5123658.1	*Enterococcus faecium* FSIS11816959	
	HAZ1055918.1	*Enterococcus faecium* E1037	
	MDB7617239.1	*Enterococcus faecium* 1001295st3_G8_1001295H_181012	
	EEU87319.1	*Enterococcus faecalis* ARO1/DG	Ent-TB
	NSR64079.1	*Enterococcus faecalis* HC_NS0170	
	NTJ92680.1	*Enterococcus faecium* SWEntR-0365	
	UTJ09864.1	*Enterococcus faecalis* Chr-JH 2-2	
	NTP75778.1	*Enterococcus faecium* HC_NS0120	
	NTL63874.1	*Enterococcus faecalis* L6D	
	EJC3745485.1	*Enterococcus faecium* FSIS12210887	
	MBE6171023.1	*Enterococcus casseliflavus* SIG191	
	BDQ56092.1	*Enterococcus faecalis* SVR2331	
	PNK93531.1	*Enterococcus faecalis* FDAARGOS_347	
	MBS6192045.1	*Enterococcus* hirae L3_058_072G1_dasL3_058_072G1_metabat.metabat.14	
	OQO44504.1	*Enterococcus* hirae F0221F 224	
	BDQ49338.1	*Enterococcus faecalis* SVR2281	
	EGO2586433.1	*Enterococcus faecalis* FSIS12033586	
	EGO9214630.1	*Enterococcus faecalis* FSIS11810505	
	QYI86635.1	*Enterococcus* phage SEsuP-1	Ent-TC
	BDX45315.1	*Enterococcus faecalis* L6D	
	EGO6516563.1	*Enterococcus faecalis* FSIS32003318	
	EOL19207.1	*Enterococcus faecalis* EnGen0331 Merz89	
	EOE44992.1	*Enterococcus faecalis* EnGen0067 B1290	
	BDQ56825.1	*Enterococcus faecalis* SVR2331	
	NSM73460.1	*Enterococcus faecalis* SW_0238	
	NSR43685.1	*Enterococcus faecalis* HC_SS0017	
	EHA3992932.1	*Enterococcus faecalis* FSIS12036838	
	EKK5303559.1	*Enterococcus faecalis* FSIS12318911	

	EGO8074180.1	*Enterococcus faecalis* FSIS11920162	
	NSQ62537.1	*Enterococcus faecalis* HC_SS0005	
	EHV2873836.1	*Enterococcus faecalis* FSIS12141403	
	NSO16553.1	*Enterococcus faecalis* WW_0067J	
	EHA4022804.1	*Enterococcus faecalis* FSIS12036839	

Table A.4. Integrases from temperate phages that infect *Staphylococcus* spp. collected from NCBI GenBank and their corresponding cluster based on heatmap analysis.

Recombinase family	Accession No.	Organism	Cluster
Serine	UKM36543.1	*Staphylococcus* phage vB_SauS_Mh15	Sta-SA
	YP_007236622.1	*Staphylococcus* phage StB20	
	YP_009302049.1	*Staphylococcus* phage CNPx	
	YP_950693.1	*Staphylococcus* phage PH15	
	QQV93178.1	*Staphylococcus* virus vB_SepS_27	
	QQV93448.1	*Staphylococcus* virus vB_SepS_E72	
	YP_010080142.1	*Staphylococcus* phage UPMK_2	
	YP_007236569.1	*Staphylococcus* phage StB27	
	QQV93313.1	*Staphylococcus* phage vB_SepS_456	
	YP_001604091.1	*Staphylococcus* phage phiMR11	
	YP_009130680.1	*Staphylococcus* phage StB12	
	WEU69960.1	*Staphylococcus* phage vB_SepS_BE20	
	QQV93246.1	*Staphylococcus* virus vB_SepS_48	
Tyrosine	WP_049294684.1	*Staphylococcus aureus* RP7	Sta-TA
	QUU29388.1	*Staphylococcus* phage B_UFSM1	
	AAL82229.1	*Staphylococcus* phage phi 11	
	ASU01410.1	*Staphylococcus* phage SN8	
	UKM36001.1	*Staphylococcus* phage vB_SauS_321c	
	USL86396.1	*Staphylococcus* phage Sushi	
	YP_009209308.1	*Staphylococcus* phage 3MRA	
	AKC04676.1	*Staphylococcus* phage B236	
	APD19894.1	*Staphylococcus* phage SpT252	
	YP_001004330.1	*Staphylococcus* phage phiETA3	
	YP_001004261.1	*Staphylococcus* phage phiETA2	
	ABF73031.1	*Staphylococcus* phage phinm1	
	APD19759.1	*Staphylococcus* phage SpT5	
	AKC04612.1	*Staphylococcus* phage B166	
	UYD72034.1	*Staphylococcus* phage PHB40a	
	NP_510895.1	*Staphylococcus* phage phiETA	
	UKM35799.1	*Staphylococcus* phage vB_SauS_287	
	YP_240491.1	*Staphylococcus* phage 55	

	UUV45122.1	*Staphylococcus* phage SAP1_TA-2022	
	QWT56486.1	*Staphylococcus* phage DSP07	
	UUV45259.1	*Staphylococcus* phage SAP3_TA-2022	
	APW79902.1	*Staphylococcus* phage 3 AJ-2017	Sta-TB
	AFN39870.1	*Staphylococcus* phage StauST398-1	
	YP_009196767.1	*Staphylococcus* phage 23MRA	
	AXF39071.1	*Staphylococcus* phage SA345ruMSSAST8	
	AJT61506.1	*Staphylococcus* phage IME-SA4	
	AAL82331.1	*Staphylococcus* phage phi 13	
	AZB66813.1	*Staphylococcus* phage phiSP119-3	
	AXY85989.1	*Staphylococcus* phage SA780ruMSSAST101	
	ARM68470.1	*Staphylococcus* phage IME1365_01	
	ARM68193.1	*Staphylococcus* phage IME1367_01	
	NP_835517.1	*Staphylococcus* phage phiN315	
	YP_239821.1	*Staphylococcus* phage 2638A	
	gb|QQV87988.1	*Staphylococcus* phage PMBT9	
	ABF73161.1	*Staphylococcus* phage phiNM3	
	WGL32330.1	*Staphylococcus* phage phiST9-C	
	YP_239890.1	*Staphylococcus* phage 42E	
	QLF86510.1	*Staphylococcus* phage vB_SepS_BE01	
	AFM73733.1	*Staphylococcus* phage vB_SepiS-phiIPLA5	
	QAU03487.1	*Staphylococcus* phage Henu2	
	ARM68333.1	*Staphylococcus* phage IME1354_01	
	ARM67749.1	*Staphylococcus* phage IME1318_01	
	UUV45325.1	*Staphylococcus* phage SAP8_Andrews	
	UKM36468.1	*Staphylococcus* phage vB_SauS_Mh1	
	QXV71952.1	*Staphylococcus* phage vB_SarS_BM31	
	ANM47008.1	*Staphylococcus* phage vB_SauS_IMEP5	
	WEU70768.1	*Staphylococcus* phage vB_SepS_BE28	
	AAL82282.1	*Staphylococcus* phage phi 12	Sta-TC
	YP_009188673.1	*Staphylococcus* phage phiJB	
	YP_007237130	*Staphylococcus* phage phi5967PVL	
	YP_007005716.1	*Staphylococcus* phage SMSAP5	
	AUM58294.1	*Staphylococcus* phage phiSa2wa_st121mssa	
	AGG36598.1	*Staphylococcus* phage LH1	

	AGZ17439.1	*Staphylococcus* phage YMC/09/04/R1988	
	QAY02689.1	*Staphylococcus* phage vB_SauS_fPfSau02	
	UVD41663.1	*Staphylococcus* phage A12-4PHSA25	
	NP_958623.1	*Staphylococcus* phage 77	
	YP_009268644.1	*Staphylococcus* phage 80	
	YP_002332364.1	*Staphylococcus* phage phiSauS-IPLA35	

Table A.5. Identity of environmental phage isolates based on whole genomes and major capsid proteins.

Host bacterium	Phage isolate ID	Identity	
		Whole genome	**Major capsid protein**
E. faecalis 8413	E9-1	*Enterococcus* phage vB_EfaS_Ef5.4	*Enterococcus* phage vB_EfaS_LM99
	E23-1	*Enterococcus* phage vB_EfaS_Ef5.4	*Enterococcus* phage PMBT2
	E23-2	*Enterococcus* phage vB_EfaS_Ef5.4	*Enterococcus* phage PMBT2
	E49-2	*Enterococcus* phage vB_EfaS_Ef5.4	*Enterococcus* phage vB_EfaS_LM99
	E51-2	*Enterococcus* phage vB_EfaS_Ef5.4	*Enterococcus* phage vB_EfaS_LM99
	E102-1	*Enterococcus* phage vB_EFaS_TV51	*Enterococcus* phage vB_Efa_VP16
	E102-2	*Enterococcus* phage vB_EFaS_TV51	*Enterococcus* phage vB_Efa_VP16
S. aureus RP29	SA2-1	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	SA28-2	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	SA48-2	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	SA50-2	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	SA52-2	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	CIP93	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	CIP94	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome
	CIP95	*Staphylococcus* phage Portland	*Staphylococcus* phage vB_SauP-436A1
	CIP98	*Staphylococcus aureus* pt202 chromosome	*Staphylococcus aureus* pt202 chromosome

References

1. Ventola CL. The Antibiotic Resistance Crisis: Part 1: Causes and Threats. P T Peer-Rev J Formul Manag. 2015 Apr;40(4):277–83.

2. Mancuso G, Midiri A, Gerace E, Biondo C. Bacterial Antibiotic Resistance: The Most Critical Pathogens. Pathog Basel Switz. 2021 Oct 12;10(10):1310.

3. Llor C, Bjerrum L. Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem. Ther Adv Drug Saf. 2014 Dec;5(6):229–41.

4. WHO Bacterial Priority Pathogens List 2024: Bacterial Pathogens of Public Health Importance, to Guide Research, Development, and Strategies to Prevent and Control Antimicrobial Resistance. 1st ed. Geneva: World Health Organization; 2024. 1 p.

5. Asokan GV, Ramadhan T, Ahmed E, Sanad H. WHO Global Priority Pathogens List: A Bibliometric Analysis of Medline-PubMed for Knowledge Mobilization to Infection Prevention and Control Practices in Bahrain. Oman Med J. 2019 May;34(3):184–93.

6. Schwartzman JA, Lebreton F, Salamzade R, Shea T, Martin MJ, Schaufler K, et al. Global diversity of enterococci and description of 18 previously unknown species. Proc Natl Acad Sci. 2024 Mar 5;121(10):e2310852121.

7. Kalode V, Patil P. Enterococcus Species: A Systemic Review. J Pure Appl Microbiol. 2023 Jun 1;17(2):761–7.

8. Lebreton F, Willems RJL, Gilmore MS. Enterococcus Diversity, Origins in Nature, and Gut Colonization. In: Gilmore MS, Clewell DB, Ike Y, Shankar N, editors. Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet]. Boston: Massachusetts Eye and Ear Infirmary; 2014 [cited 2025 Feb 28]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK190427/

9. Hammerum AM. Enterococci of animal origin and their significance for public health. Clin Microbiol Infect. 2012 Jul;18(7):619–25.

10. Ramos S, Silva V, Dapkevicius M, Igrejas G, Poeta P. Enterococci, from Harmless Bacteria to a Pathogen. Microorganisms. 2020 Jul 25;8(8):1118.

11. García-Solache M, Rice LB. The Enterococcus: a Model of Adaptability to Its Environment. Clin Microbiol Rev. 2019 Mar 20;32(2):e00058-18.

12. Hollenbeck BL, Rice LB. Intrinsic and acquired resistance mechanisms in enterococcus. Virulence. 2012 Aug 15;3(5):421–33.

13. Joshi S, Shallal A, Zervos M. Vancomycin-Resistant Enterococci. Infect Dis Clin North Am. 2021 Dec;35(4):953–68.

14. Sattari-Maraji A, Jabalameli F, Node Farahani N, Beigverdi R, Emaneini M. Antimicrobial resistance pattern, virulence determinants and molecular analysis of Enterococcus faecium isolated from children infections in Iran. BMC Microbiol. 2019 Dec;19(1):156.

15. Fiore E, Van Tyne D, Gilmore MS. Pathogenicity of Enterococci. Fischetti VA, Novick RP, Ferretti JJ, Portnoy DA, Braunstein M, Rood JI, editors. Microbiol Spectr. 2019 Jul 19;7(4):7.4.9.

16. Dadashi M, Sharifian P, Bostanshirin N, Hajikhani B, Bostanghadiri N, Khosravi-Dehaghi N, et al. The Global Prevalence of Daptomycin, Tigecycline, and Linezolid-Resistant Enterococcus faecalis and Enterococcus faecium Strains From Human Clinical Samples: A Systematic Review and Meta-Analysis. Front Med. 2021;8:720647.

17. Bender JK, Cattoir V, Hegstad K, Sadowy E, Coque TM, Westh H, et al. Update on prevalence and mechanisms of resistance to linezolid, tigecycline and daptomycin in enterococci in Europe: Towards a common nomenclature. Drug Resist Updat. 2018 Sep;40:25–39.

18. Chenoweth C, Schaberg D. The epidemiology of enterococci. Eur J Clin Microbiol Infect Dis. 1990 Feb;9(2):80–9.

19. Moghimbeigi A, Moghimbeygi M, Dousti M, Kiani F, Sayehmiri F, Sadeghifard N, et al. Prevalence of vancomycin resistance among isolates of enterococci in Iran: a systematic review and meta-analysis. Adolesc Health Med Ther. 2018;9:177–88.

20. Boccella M, Santella B, Pagliano P, De Filippis A, Casolaro V, Galdiero M, et al. Prevalence and Antimicrobial Resistance of Enterococcus Species: A Retrospective Cohort Study in Italy. Antibiot Basel Switz. 2021 Dec 19;10(12):1552.

21. Gordon S, Swenson JM, Hill BC, Pigott NE, Facklam RR, Cooksey RC, et al. Antimicrobial susceptibility patterns of common and unusual species of enterococci causing infections in the United States. Enterococcal Study Group. J Clin Microbiol. 1992 Sep;30(9):2373–8.

22. Khan A, Miller WR, Axell-House D, Munita JM, Arias CA. Antimicrobial Susceptibility Testing for Enterococci. J Clin Microbiol. 2022 Sep 21;60(9):e0084321.

23. Krawczyk B, Wityk P, Gałęcka M, Michalik M. The Many Faces of Enterococcus spp.—Commensal, Probiotic and Opportunistic Pathogen. Microorganisms. 2021 Sep 7;9(9):1900.

24. Gillespie SH. Gram-positive cocci. In: Medical Microbiology Illustrated [Internet]. Elsevier; 1994 [cited 2025 Mar 3]. p. 12–29. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780750601870500079

25. Hatoum-Aslan A. The phages of staphylococci: critical catalysts in health and disease. Trends Microbiol. 2021 Dec;29(12):1117–29.

26. Bonar E, Międzobrodzki J, Władyka B. The Staphylococcal Coagulases. In: Pet-To-Man Travelling Staphylococci [Internet]. Elsevier; 2018 [cited 2025 Mar 17]. p. 95–102. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128135471000078

27. Hennekinne JA, Le Loir Y. STAPHYLOCOCCUS | Detection by Cultural and Modern Techniques. In: Encyclopedia of Food Microbiology [Internet]. Elsevier; 2014 [cited 2025 Mar 3]. p. 487–93. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123847300003189

28. Cangui-Panchi SP, Ñacato-Toapanta AL, Enríquez-Martínez LJ, Salinas-Delgado GA, Reyes J, Garzon-Chavez D, et al. Battle royale: Immune response on biofilms - host-pathogen interactions. Curr Res Immunol. 2023;4:100057.

29. Guo Y, Song G, Sun M, Wang J, Wang Y. Prevalence and Therapies of Antibiotic-Resistance in Staphylococcus aureus. Front Cell Infect Microbiol. 2020 Mar 17;10:107.

30. Gajdács M. The Continuing Threat of Methicillin-Resistant Staphylococcus aureus. Antibiot Basel Switz. 2019 May 2;8(2):52.

31. Kaur DC, Chate SS. Study of Antibiotic Resistance Pattern in Methicillin Resistant Staphylococcus Aureus with Special Reference to Newer Antibiotic. J Glob Infect Dis. 2015;7(2):78–84.

32. Appelbaum PC. MRSA—the tip of the iceberg. Clin Microbiol Infect. 2006;12:3–10.

33. Chiang HY, Perencevich EN, Nair R, Nelson RE, Samore M, Khader K, et al. Incidence and Outcomes Associated With Infections Caused by Vancomycin-Resistant Enterococci in the United States: Systematic Literature Review and Meta-Analysis. Infect Control Hosp Epidemiol. 2017 Feb;38(2):203–15.

34. McGuinness WA, Malachowa N, DeLeo FR. Vancomycin Resistance in Staphylococcus aureus. Yale J Biol Med. 2017 Jun;90(2):269–81.

35. Sharma NK, Garg R, Baliga S, Bhat K G. Nosocomial Infections and Drug Susceptibility Patterns in Methicillin Sensitive and Methicillin Resistant Staphylococcus aureus. J Clin Diagn Res JCDR. 2013 Oct;7(10):2178–80.

36. Sollid JUE, Furberg AS, Hanssen AM, Johannessen M. Staphylococcus aureus: Determinants of human carriage. Infect Genet Evol. 2014 Jan;21:531–41.

37. Laux C, Peschel A, Krismer B. Staphylococcus aureus Colonization of the Human Nose and Interaction with Other Microbiome Members. Microbiol Spectr. 2019 Mar;7(2).

38. Lawrence RM. Transmission of Infectious Diseases Through Breast Milk and Breastfeeding. In: Breastfeeding [Internet]. Elsevier; 2011 [cited 2025 Mar 4]. p. 406–73. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9781437707885100136

39. Gurung RR, Maharjan P, Chhetri GG. Antibiotic Resistance Pattern of *Staphylococcus Aureus* With Reference to MRSA Isolates From Pediatric Patients. Future Sci OA. 2020 Apr;6(4):FSO464.

40. Tălăpan D, Sandu AM, Rafila A. Antimicrobial Resistance of Staphylococcus aureus Isolated between 2017 and 2022 from Infections at a Tertiary Care Hospital in Romania. Antibiotics. 2023 May 28;12(6):974.

41. Dong Y, Speer CP. The role of Staphylococcus epidermidis in neonatal sepsis: Guarding angel or pathogenic devil? Int J Med Microbiol. 2014 Jul;304(5–6):513–20.

42. Rogers KL, Fey PD, Rupp ME. Coagulase-negative staphylococcal infections. Infect Dis Clin North Am. 2009 Mar;23(1):73–98.

43. Hibstu Z, Belew H, Akelew Y, Mengist HM. Phage Therapy: A Different Approach to Fight Bacterial Infections. Biol Targets Ther. 2022 Oct;Volume 16:173–86.

44. Harding KR, Kyte N, Fineran PC. Jumbo phages. Curr Biol CB. 2023 Jul 24;33(14):R750–1.

45. Strathdee SA, Hatfull GF, Mutalik VK, Schooley RT. Phage therapy: From biological mechanisms to future directions. Cell. 2023 Jan 5;186(1):17–31.

46. Fong K, Wong CWY, Wang S, Delaquis P. How Broad Is Enough: The Host Range of Bacteriophages and Its Impact on the Agri-Food Sector. PHAGE New Rochelle N. 2021 Jun 1;2(2):83–91.

47. Dwivedi B, Schmieder R, Goldsmith DB, Edwards RA, Breitbart M. PhiSiGns: an online tool to identify signature genes in phages and design PCR primers for examining phage diversity. BMC Bioinformatics. 2012 Mar 4;13:37.

48. Hatfull GF, Hendrix RW. Bacteriophages and their genomes. Curr Opin Virol. 2011 Oct;1(4):298–303.

49. Al-Shayeb B, Sachdeva R, Chen LX, Ward F, Munk P, Devoto A, et al. Clades of huge phages from across Earth's ecosystems. Nature. 2020 Feb 20;578(7795):425–31.

50. Moura de Sousa JA, Pfeifer E, Touchon M, Rocha EPC. Causes and Consequences of Bacteriophage Diversification via Genetic Exchanges across Lifestyles and Bacterial Taxa. Arkhipova I, editor. Mol Biol Evol. 2021 May 19;38(6):2497–512.

51. Smug BJ, Szczepaniak K, Rocha EPC, Dunin-Horkawicz S, Mostowy RJ. Ongoing shuffling of protein fragments diversifies core viral functions linked to interactions with bacterial hosts. Nat Commun. 2023 Nov 28;14(1):7460.

52. Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev. 2014 Sep;38(5):916–31.

53. Lima-Mendez G, Van Helden J, Toussaint A, Leplae R. Reticulate Representation of Evolutionary and Functional Relationships between Phage Genomes. Mol Biol Evol. 2008 Feb 14;25(4):762–77.

54. Sinha S, Grewal RK, Roy S. Modeling Bacteria–Phage Interactions and Its Implications for Phage Therapy. In: Advances in Applied Microbiology [Internet]. Elsevier; 2018 [cited 2024 Jun 19]. p. 103–41. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065216418300054

55. Stone E, Campbell K, Grant I, McAuliffe O. Understanding and Exploiting Phage-Host Interactions. Viruses. 2019 Jun 18;11(6):567.

56. Kropinski AM. Phage Therapy -- Everything Old Is New again. Can J Infect Dis Med Microbiol. 2006;17(5):297–306.

57. Zhang M, Zhang T, Yu M, Chen YL, Jin M. The Life Cycle Transitions of Temperate Phages: Regulating Factors and Potential Ecological Implications. Viruses. 2022 Aug 28;14(9):1904.

58. Groth AC, Calos MP. Phage Integrases: Biology and Applications. J Mol Biol. 2004 Jan;335(3):667–78.

59. Feiner R, Argov T, Rabinovich L, Sigal N, Borovok I, Herskovits AA. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. Nat Rev Microbiol. 2015 Oct;13(10):641–50.

60. Wahl LM, Pattenden T. Prophage Provide a Safe Haven for Adaptive Exploration in Temperate Viruses. Genetics. 2017 May;206(1):407–16.

61. Nanda AM, Thormann K, Frunzke J. Impact of Spontaneous Prophage Induction on the Fitness of Bacterial Populations and Host-Microbe Interactions. Margolin W, editor. J Bacteriol. 2015 Feb 1;197(3):410–9.

62. Fogg PCM, Colloms S, Rosser S, Stark M, Smith MCM. New Applications for Phage Integrases. J Mol Biol. 2014 Jul;426(15):2703–16.

63. Argos P, Landy A, Abremski K, Egan JB, Haggard-Ljungquist E, Hoess RH, et al. The integrase family of site-specific recombinases: regional similarities and global diversity. EMBO J. 1986 Feb;5(2):433–40.

64. Balding C, Bromley SA, Pickup RW, Saunders JR. Diversity of phage integrases in Enterobacteriaceae: development of markers for environmental analysis of temperate phages. Environ Microbiol. 2005 Oct;7(10):1558–67.

65. Viertel TM, Ritter K, Horz HP. Viruses versus bacteria--novel approaches to phage therapy as a tool against multidrug-resistant pathogens. J Antimicrob Chemother. 2014 Sep 1;69(9):2326–36.

66. Gordillo Altamirano FL, Barr JJ. Phage Therapy in the Postantibiotic Era. Clin Microbiol Rev. 2019 Mar 20;32(2):e00066-18.

67. Chang C, Yu X, Guo W, Guo C, Guo X, Li Q, et al. Bacteriophage-Mediated Control of Biofilm: A Promising New Dawn for the Future. Front Microbiol. 2022 Apr 4;13:825828.

68. Ferriol-González C, Domingo-Calap P. Phages for Biofilm Removal. Antibiot Basel Switz. 2020 May 21;9(5):268.

69. Kovacs CJ, Rapp EM, McKenzie SM, Mazur MZ, Mchale RP, Brasko B, et al. Disruption of Biofilm by Bacteriophages in Clinically Relevant Settings. Mil Med. 2024 May 18;189(5–6):e1294–302.

70. Kifelew LG, Warner MS, Morales S, Thomas N, Gordon DL, Mitchell JG, et al. Efficacy of Lytic Phage Cocktails on Staphylococcus aureus and Pseudomonas aeruginosa in Mixed-Species Planktonic Cultures and Biofilms. Viruses. 2020 May 18;12(5):559.

71. Cui Z, Guo X, Feng T, Li L. Exploring the whole standard operating procedure for phage therapy in clinical practice. J Transl Med. 2019 Dec;17(1):373.

72. Gill J, Hyman P. Phage Choice, Isolation, and Preparation for Phage Therapy. Curr Pharm Biotechnol. 2010 Jan 1;11(1):2–14.

73. Olsen NS, Hendriksen NB, Hansen LH, Kot W. A New High-Throughput Screening Method for Phages: Enabling Crude Isolation and Fast Identification of Diverse Phages with Therapeutic Potential. PHAGE. 2020 Sep 1;1(3):137–48.

74. Kenney PO, Gómez-Duarte OG. Low-volume enrichment method supports high throughput bacteriophage screening and isolation from wastewater. Chang YF, editor. PLOS ONE. 2024 Apr 16;19(4):e0298833.

75. Doss JH, Barekzi N, Gauthier DT. Improving high-throughput techniques for bacteriophage discovery in multi-well plates. J Microbiol Methods. 2022 Sep;200:106542.

76. Aslam S, Lampley E, Wooten D, Karris M, Benson C, Strathdee S, et al. Lessons Learned From the First 10 Consecutive Cases of Intravenous Bacteriophage Therapy to Treat Multidrug-Resistant Bacterial Infections at a Single Center in the United States. Open Forum Infect Dis. 2020 Sep 1;7(9):ofaa389.

77. Dedrick RM, Guerrero-Bustamante CA, Garlena RA, Russell DA, Ford K, Harris K, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant Mycobacterium abscessus. Nat Med. 2019 May;25(5):730–3.

78. Kauffman KM, Polz MF. Streamlining standard bacteriophage methods for higher throughput. MethodsX. 2018;5:159–72.

79. Hyman P. Phages for Phage Therapy: Isolation, Characterization, and Host Range Breadth. Pharmaceuticals. 2019 Mar 11;12(1):35.

80. Mutalik VK, Arkin AP. A Phage Foundry Framework to Systematically Develop Viral Countermeasures to Combat Antibiotic-Resistant Bacterial Pathogens. iScience. 2022 Apr;25(4):104121.

81. Doub JB. Risk of Bacteriophage Therapeutics to Transfer Genetic Material and Contain Contaminants Beyond Endotoxins with Clinically Relevant Mitigation Strategies. Infect Drug Resist. 2021 Dec;Volume 14:5629–37.

82. Born Y, Knecht LE, Eigenmann M, Bolliger M, Klumpp J, Fieseler L. A major-capsid-protein-based multiplex PCR assay for rapid identification of selected virulent bacteriophage types. Arch Virol. 2019 Mar;164(3):819–30.

83. Smith KC, Castro-Nallar E, Fisher JN, Breakwell DP, Grose JH, Burnett SH. Phage cluster relationships identified through single gene analysis. BMC Genomics. 2013 Dec;14(1):410.

84. Kornienko M, Bespiatykh D, Malakhova M, Gorodnichev R, Kuptsov N, Shitikov E. PCR Assay for Rapid Taxonomic Differentiation of Virulent Staphylococcus aureus and Klebsiella pneumoniae Bacteriophages. Int J Mol Sci. 2023 Feb 24;24(5):4483.

85. Colavecchio A, D'Souza Y, Tompkins E, Jeukens J, Freschi L, Emond-Rheault JG, et al. Prophage Integrase Typing Is a Useful Indicator of Genomic Diversity in Salmonella enterica. Front Microbiol. 2017 Jul 10;8:1283.

86. Kropinski AM, Mazzocco A, Waddell TE, Lingohr E, Johnson RP. Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay. In: Clokie MRJ, Kropinski AM, editors. Bacteriophages [Internet]. Totowa, NJ: Humana Press; 2009 [cited 2025 Mar 14]. p. 69–76. (Walker JM, editor. Methods in Molecular Biology; vol. 501). Available from: http://link.springer.com/10.1007/978-1-60327-164-6_7

87. Serwer P, Hayes SJ, Thomas JA, Demeler B, Hardies SC. Isolation of Novel Large and Aggregating Bacteriophages. In: Clokie MRJ, Kropinski AM, editors. Bacteriophages [Internet]. Totowa, NJ: Humana Press; 2009 [cited 2025 Mar 14]. p. 55–66. (Walker JM, editor. Methods in Molecular Biology; vol. 501). Available from: http://link.springer.com/10.1007/978-1-60327-164-6_6

88. Bonilla N, Barr JJ. Phage on Tap: A Quick and Efficient Protocol for the Preparation of Bacteriophage Laboratory Stocks. In: Moya A, Pérez Brocal V, editors. The Human Virome [Internet]. New York, NY: Springer New York; 2018 [cited 2025 Mar 14]. p. 37–46. (Methods in Molecular Biology; vol. 1838). Available from: http://link.springer.com/10.1007/978-1-4939-8682-8_4

89. Pickard DJJ. Preparation of Bacteriophage Lysates and Pure DNA. In: Clokie MRJ, Kropinski AM, editors. Bacteriophages [Internet]. Totowa, NJ: Humana Press; 2009 [cited 2025 Mar 19]. p. 3–9. (Walker JM, editor. Methods in Molecular Biology; vol. 502). Available from: http://link.springer.com/10.1007/978-1-60327-565-1_1

90. Brives C, Pourraz J. Phage therapy as a potential solution in the fight against AMR: obstacles and possible futures. Palgrave Commun. 2020 May 19;6(1):100.

91. Suh GA, Lodise TP, Tamma PD, Knisely JM, Alexander J, Aslam S, et al. Considerations for the Use of Phage Therapy in Clinical Practice. Antimicrob Agents Chemother. 2022 Mar 15;66(3):e02071-21.

92. Grose JH, Casjens SR. Understanding the enormous diversity of bacteriophages: The tailed phages that infect the bacterial family Enterobacteriaceae. Virology. 2014 Nov;468–470:421–43.

93. Dini M, Shokoohizadeh L, Jalilian FA, Moradi A, Arabestani MR. Genotyping and characterization of prophage patterns in clinical isolates of Staphylococcus aureus. BMC Res Notes. 2019 Oct 21;12(1):669.

94. Rostami K, Nemati M, Pourahmad F. Prophage Typing of Staphylococcus aureus Strains Isolated from Broiler Poultry. Arch Razi Inst. 2021;76(3):499–506.

95. Goerke C, Pantucek R, Holtfreter S, Schulte B, Zink M, Grumann D, et al. Diversity of Prophages in Dominant *Staphylococcus aureus* Clonal Lineages. J Bacteriol. 2009 Jun;191(11):3462–8.

96. Wolput S, Lood C, Fillol-Salom A, Casters Y, Albasiony A, Cenens W, et al. Phage-host co-evolution has led to distinct generalized transduction strategies. Nucleic Acids Res. 2024 Jul 22;52(13):7780–91.

97. Parcey M, Gayder S, Castle AJ, Svircev AM. Molecular Profile of Phage Infection: A Novel Approach for the Characterization of Erwinia Phages through qPCR. Int J Mol Sci. 2020 Jan 15;21(2):553.

98. Bichet MC, Patwa R, Barr JJ. Protocols for studying bacteriophage interactions with in vitro epithelial cell layers. STAR Protoc. 2021 Sep 17;2(3):100697.

99. Sobolewski I, Adamowicz K, Struck A, Zylicz-Stachula A, Skowron PM. A Method for Isolation Bacteriophage Particles-Free Genomic DNA, Exemplified by TP-84, Infecting Thermophilic Geobacillus. Microorganisms. 2022 Sep 3;10(9):1782.

100. Kornienko M, Bespiatykh D, Abdraimova N, Gorodnichev R, Gostev V, Boldyreva D, et al. Multiomics analysis of Staphylococcus aureus ST239 strains resistant to virulent Herelleviridae phages. Sci Rep. 2024 Nov 26;14(1):29375.