Smarcs vs *Serratia:* Characterising Novel Phages to Combat *Serratia marcescens*

Beau Jack Patrick

Bachelor of Science

A thesis submitted in total fulfilment of the requirements for the degree of Master of Science

Department of Physiology, Anatomy and Microbiology
College of Science, Health and Engineering
La Trobe University, Victoria, Australia

December 2021

Table of Contents

List of Tables

List of Figures

Statement of Authorship

I certify that the attached document is my original work. No other person's work has been used without due acknowledgement in the text of the document.

Except where reference is made in the text, this document contains no material presented elsewhere or extracted in whole or in part from a document presented by me for another qualification at this or another institution.

None of the research undertaken in connection with this thesis required approval by a University Ethics Committee.

Beau J. Patrick
23rd December 2021

Acknowledgements

I would like to say a massive thank you to Dr Steve Petrovski, my supervisor, for giving me the opportunity to do my honours and masters at the Petrovski Lab. The last two years presented some unexpected challenges. Despite this, however, you made it a fun experience and helped each of the lab members to make the most of a difficult situation. I am proud of the work we have achieved together and never could have done it without your advice and encouragement. This past year has been full of laughs, lessons, and experiences that I will never forget, thank you.

I'd also like to thank my co-supervisor, Dr Mark Chan. I really value your insights and appreciate the time you took to explain ideas to me in detail. You helped a lot with my understanding of concepts and provided me with a broader perspective of things. Thank you for the opportunity to do some extra work during the year, it was a great experience and I look forward to hearing about your involvement in the start-up community in the future.

Thank you to the members of the Petrovski lab who made our time there special: Aurelie, Cassie, Jay, Kelsey, Damian, Vaheesan and Steven. It was wonderful becoming friends with you all and I thank you for the support and fun times that we had. A special thank you to Aurelie, Cassie, Jay and Vahee for taking the time to share your knowledge with me while I learned the ropes. I'd also like to mention our neighbours Lauren, Ella and Will; the constant chats and laughs always brightened my day.

Finally, thank you to my amazing friends and family. Your enthusiasm and support throughout this journey have been invaluable. Thank you, Mum & Dad, I couldn't have done it without you.

This work was supported by an Australian Government Research Training Program Scholarship.

Conference Presentations

2021 **Australian Society of Microbiology National Conference.** Virtual poster presentation. “Network Phylogeny: *Serratia* Phages in a Sea of Sequences

2021 **Australian Society of Microbiology Nancy Millis Student Awards Night.** Awarded the Nancy Millis Vic Branch award for a 10-minute presentation entitled “Network Phylogeny: *Serratia* Phages in a Sea of Sequences”.

List of Abbreviations

α	Alpha
bp	Base pair (s)
°C	Degrees Celsius
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid disodium salt dihydrate
et al	*Et alii* (and others)
g	Gram (s)
x g	Gravitational force
hr	Hour (s)
kbp	Kilo base pair (s)
λ	Lambda
kV	Kilo volt (s)
L	Litre
µg	Microgram (s)
µm	Micrometre (s)
µl	Microlitre (s)
M	Molar
MDR	Multi drug resistant
mg	Milligram (s)
min	Minute (s)
ml	Millilitre (s)
mM	Millimolar
NA	Nutrient agar
NB	Nutrient broth
NCBI	National Centre for Biotechnology Information
ng	Nanogram (s)
Nm	Nanometre (s)
Ω	Ohm (s) (resistance)
ORF	Open reading frame
PCR	Polymerase chain reaction
PEG 8000	Polyethylene glycol 8000

$^{-1}$	per
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
sec	Second (s)
S.O.C.	Super optimal culture
TAE	Tris-acetate-EDTA
U	Units
V	Volts
v/v	Volume per volume
w/v	Weight per volume
WGS	Whole genome sequencing

COVID impact statement

Dear Thesis Examiner,

In response to the COVID-19 pandemic, from March 2020 – February 2021, I did not have laboratory access. This had severe impacts on my work as my project is mostly lab-based. Before lockdown was imposed in Melbourne, we had already begun work towards the *in vitro* characterisation of 5 novel bacteriophages, however my project had to be completely reimagined so that it could be worked on from home. During the prolonged lockdown in 2020, I performed the bioinformatic analysis detailed herein. When lab access was permitted in 2021 the *in vitro* characterisation of the phages was able to recommence, albeit with a heavily condensed timeline. As such, I was unable to complete some of the work that I was hoping to accomplish before submitting my thesis.

Please consider these factors when assessing this thesis.

Thank you for your understanding,
Beau J. Patrick

Summary

Serratia marcescens is a Gram-negative bacterium that belongs to a group of pathogenic organisms impacting health and agriculture. These organisms are multi-drug resistant pathogens and threaten the health of humans and animals. The emergence of multi-drug resistant *S. marcescens* has arisen from the overuse of antibiotics in medical treatment and industrial applications. *S. marcescens* harbours an arsenal of virulence factors contributing to its pathogenicity and targets hospitalised and immunocompromised hosts. Due to an array of intrinsic and acquired antibiotic resistance determinants hosted by *S. marcescens*, treating infections with traditional antibiotic therapies is becoming increasingly problematic. In addition, negative health implications associated with antibiotic exposure in humans and animals are a growing concern. These issues are promoting interest in treatments to augment or replace antibiotics.

Phage therapy is the use of bacteriophages (phages) in the treatment of bacterial infections. Phages are viruses that specifically infect their bacterial hosts, meaning their exposure to humans and other eukaryotic organisms poses minimal risk. The potential medical and ecological applications of phage therapy appear promising in a world threatened by the diminishing effectiveness of our current antibiotics. With the vastness of their genetic diversity becoming increasingly apparent, efforts to give phages predictive and explanatory classifications are proving invaluable. Herein, six novel *S. marcescens* phages are introduced and consolidated into clusters of related genomes. In addition, through gene content analysis, a reticulated network phylogeny is produced that places the phages in context among the global phage sequence space. Subsequent *in vitro* characterisation is undertaken to assess the phage's therapeutic potential when applied to clinical *S. marcescens* strains individually and in phage cocktails.

Chapter 1: Literature Review

1.1 Introduction

Occurrences of multi-drug resistant (MDR) pathogens are on the rise, with some members of the scientific community believing the post-antibiotic era has dawned (1). Since the discovery of penicillin in 1928, complacent administration of antibiotics in medicine and industry has led to the development of resistant microbes, creating unforeseen implications in health and economy (2). Driven by a need for treatments to replace or augment antibiotics, interest in modes of microbial resistance and alternative therapies has risen (3). One pathogenic organism of interest is *Serratia marcescens.* This organism can cause disease both in humans and animals and will be the focus of this review.

S. marcescens has been categorised by the World Health Organisation (WHO) as a critical priority pathogen in need of new antibiotics, contributing to the growing list of pathogens threatening the immunocompromised (4). Infection due to *S. marcescens* tends to be opportunistic and most often occurs in hospitalised patients, especially those having previously been treated with antibiotics (5). Outside of human infection, *S. marcescens* is also reported to be pathogenic to insect hosts, being implicated in diseases of captive populations of honey bees (6, 7). Honey bees provide an economically and environmentally important pollinating service, with $360 billion worth of crops relying on this pollination globally (8). As such, increases in winter-loss and colony collapse disorder are an alarming agricultural issue (9). No single agent has been shown to be solely responsible (9), however, a number of studies have revealed mechanisms by which *S. marcescens* may be contributing to winter-loss and reduced honey bee fitness (10-12). Due to its implication in health and agriculture, and its intrinsic and acquired resistance to a range of antibiotics, alternative antimicrobials against *S. marcescens* are desirable.

One alternative to antibiotic treatment of *S. marcescens* infection is the utilisation of phage therapy. Phage therapy is the use of bacterial viruses, or bacteriophages (phages), in treating bacterial infections in humans, animals or plants (13). Clinical and industrial applications of phages have been demonstrated, and interest surrounding their effectiveness in different applications is growing (14, 15). An advantage for phage therapy as an alternative to antibiotics

is the knowledge that phages are not affected by host antibiotic resistance mechanisms. Increased interest in phage therapy and the growing efficiency among genomic sequencing technologies has led to a deluge of novel phages being sequenced. Driven by a desire to investigate the genetic relationships between groups of phages, contemporary genome comparison and phylogenetic approaches continue to be developed.

1.2 *Serratia marcescens*

S. marcescens is a Gram-negative, facultatively anaerobic, rod-shaped bacterium that often displays blood-red pigmentation (5). Recent changes to the taxonomic structure of the *Enterobacteriales* order has led to the reclassification of the *Serratia* genus from the *Enterobacteriaceae* family to the *Yersiniaceae* family. As a result, all pre-2016 publications classify the *Serratia* genus as a member of the *Enterobacteriaceae* family (16).

Different strains of *S. marcescens* have been isolated from a range of different hosts and environments such as hospitalised patients, small mammals, insects, plants, soil and water (17). It is preferably anaerobic, however in nutrient rich environments it can grow aerobically and semi-anaerobically (5). It will grow anaerobically even in nutrient sparse environments such as deionised water. Its persistence in nutrient limited conditions is highlighted by its isolation from disinfectants and antiseptics (18, 19), as well as double distilled water (20). This hardiness is attributed to its ability to utilise a wide range of nutrients (5). Its brilliant red pigmentation led to the mention of *S. marcescens* in some of the earliest works on microorganisms (21).

1.2.1 History and identification of *S. marcescens*

The prevalence of *Serratia* spp. is rooted in the annals of history. In ancient times, mysterious red droplets appearing on sacrificial bread unsettled the superstitious (22). In 1823, Italian chemist Bartolomeo Bizio coined *Serratia marcescens* upon discovering the cause of the unholy manifestation to be a microorganism (21). Initially thought to be a harmless saprophyte due to its predilection for starchy foods, it now represents a major cause of nosocomial infections (5). The first documented event involving *S. marcescens* as an aetiological agent was in 1953 involving an outbreak of 11 cases at Stanford University Hospital (23). Since then, it has gained notoriety for being a hardy pathogen involved in myriad diseases including pneumonia (24), lower respiratory tract infections (25), urinary tract infections (UTI) (26), septicaemia, wound infections and meningitis (27). Effective identification and classification of *S. marcescens* have been important in controlling clinical outbreaks and infections (28).

Biochemical identification methods have been widely used for identification of *Serratia* spp. (5), however, incorrect identification of *Serratia liquefaciens* can occur (29, 30). Modern molecular and proteomics based approaches provide more reliable classification (5). Matrix-

assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) provides rapid, automated identification of *S. marcescens* at the species level and is the most routinely deployed system (31). Identification of *S. marcescens* is also carried out via biotyping, serotyping, ribotyping and phage typing; however efficiency and reliability of each technique varies (32). Colony pigmentation is not an effective indicator of *Serratia* spp., as the production of the pigment 'prodigiosin' is dependent on growth conditions (33). Some characteristic extracellular products associated with *S. marcescens* include the presence of proteases, nucleases, a chitinase and a lipase (34). Motility of *S. marcescens* is affected by two factors, flagellation and 'serrawettin' production. Serrawettin is an extracellular surfactant that permits swarm motility in cells independent of flagellation (35, 36). The importance of efficient identification of *S. marcescens* has become increasingly apparent, as the slew of virulence factors contributing to its clinical significance have been revealed.

1.2.3 Virulence and pathogenicity in humans and insects

Several *S. marcescens* virulence factors have been documented, which act in concert to contribute to the overall pathogenicity of the organism. One such virulence factor is hemolysin. The hemolytic properties of *S. marcescens* have been attributed to two proteins; ShlA and ShlB. ShlA attaches to the host erythrocyte membrane and causes lysis via channel formation (17). This hemolytic activity of Sh1A seems to be mediated by ShlB, with evidence suggesting that Sh1B is involved in the activation of ShlA (17). These two proteins acting together have shown cytotoxic effects against host epithelial cells and fibroblasts as well as erythrocytes (37).

Adhesion is another important virulence factor involved in bacterial pathogenicity, enabling cells to resist mechanical clearing forces and initiate disease causing processes such as toxin secretion and proliferation (38). Adhesion of *Serratia* cells to mammalian epithelial cells is enabled by three types of fimbriae; type 1, type 3 and thin fimbriae, all of which are expressed by *S. marcescens* (17). This serves *S. marcescens* as a route to establish UTIs, predominantly as nosocomial secondary infections (39).

In addition to virulence factors like toxin production and adhesion, quorum sensing is an important contributor to *S. marcescens* pathogenicity. Quorum sensing is a form of cell-to-cell communication in microbes, and is facilitated by signalling molecules called *N*-acyl-homoserine lactones (AHLs) (40). Quorum sensing is crucial for biofilm development, which

is associated with increased resistance to antibiotics and host immune systems (41). Biofilm production can be observed in many nosocomial infections, contributing to colonization on catheters, tracheostomy tubes and contact lenses (37). Increased resistance to antibiotics and host immunity associated with biofilm production adds to the difficulty of clearing a *Serratia* infection.

The pathogenicity of some *S. marcescens* depend on their virulence factors and is exemplified in a 2011 report from Pinna *et al*. (42). Keratitis is a painful inflammation of the cornea, representing the most severe complication of contact-lens use (43). Gelatinase and alkaline protease production has been suggested to play a major role in the pathogenicity of *S. marcescens* keratitis (42). In addition, resistance to contact-lens disinfectant is a documented trait of *S. marcescens* and represents a challenge in medical technology (44). The production of gelatinase and alkaline protease operate in conjunction with *S. marcescens* disinfectant resistance to contribute to the pathogenesis of the disease. Whilst the above virulence factors are described only in the context of human infection, they are implemented alongside others in the infection of insects as well.

Insect pathogenicity is reportedly associated with bacterial invasion of the circulatory fluid-containing hemocoel of the host (17). Genomic analysis of *S. marcescens* isolates colonising the hemocoel of dead honey bees by Raymann *et al.* (10), identified candidate virulence genes associated with siderophore biosynthesis and flagellar production. Siderophores are metal-chelating agents important in nutrient acquisition, responsible for the acquisition of iron to be taken up into the bacterial cell (45). In the case of honey bee infection, motility conferred by flagellation allows entry into the hemocoel via the gut of infected bees (10). Chitinase is another major virulence factor in the pathogenicity of insects, with chitin being the primary component of insect's cuticle (outer body wall) (6). In keeping with *Serratia's* opportunistic nature, perturbation of host gut microbiota through use of antibiotics and insecticides has been shown to increase susceptibility to infection (46, 47). The factors described above outline the virulence and pathogenicity of *S. marcescens* and demonstrate the importance of novel bactericides, particularly in the absence of effective antibiotics.

1.2.4 Antibiotic resistance

In addition to resistance conferred by biofilm production, *Serratia* also have additional modes of antibiotic resistance. ESCAPPM is a mnemonic that denotes several organisms, including *Serratia* spp., able to demonstrate inducible, chromosomally mediated β-lactamase activity (48). *S. marcescens* displays resistance to narrow-spectrum β-lactam antibiotics, including ampicillin and first and second generation cephalosporins (5). Resistance to β-lactam antibiotics is enabled by the chromosomal *amp*C gene (49). *amp*C encodes the AmpC β-lactamase enzyme which acts by hydrolysing the drug's β-lactam ring. Chromosomally mediated resistance in *S. marcescens* has been shown to be bolstered by plasmid-mediated extended-spectrum β-lactamases (ESBLs). Plasmid-mediated ESBLs are easily disseminated within microbial communities, contributing to the spread of antibiotic resistance through horizontal gene transfer (50).

In 2017, WHO marked the discovery of new treatments for carbapenem-resistant *Enterobacteriales* species as a critical priority. The rise of carbapenem-resistant *Enterobacteriales* species is of grave concern as carbapenem antibiotics are considered the last resort when treating MDR bacterial infections (51). Alarmingly, *S. marcescens* strains expressing plasmid and chromosomally mediated carbapenemase are being isolated from clinical outbreaks (52, 53). Whilst these instances are currently considered exceptional, they highlight the growing inadequacy of existing antibiotics.

Inherent antibiotic resistance traits of environmental *S. marcescens* strains have been found to extend to the cell wall inhibitor colistin, and protein synthesis inhibitors tetracycline and nitrofurantoin (17). Environmental isolates are typically susceptible to the aminoglycosides, quinolones, chloramphenicol and trimethoprim/sulfamethoxazole, however, clinical strains displaying resistance have been isolated for each of these antibiotics (37). Strains demonstrating multi-drug resistance tend to appear in patients previously treated with one or more of these antibiotics. For example, treatment with one or more aminoglycoside may result in selection of a hyperproducing mutant of the chromosomal enzyme AAC(6′)-Ic, which is an aminoglycoside inactivating acetyltransferase (54). An AAC(6′)-Ic-hyperproducing strain would then be free to proliferate and present challenges related to amikacin, tobramycin, netilmicin, neomycin, and kanamycin resistance (54). As an increasing number of clinical *S.*

marcescens isolates are displaying multidrug resistance, there is an urgent need for novel therapeutics effective in treating *S. marcescens* infections.

1.3 Bacteriophages and phage therapy

Phages are described as obligate intracellular bacterial viruses. Comparatively smaller than their bacterial hosts, phages have genomes ranging from a few thousand to ~400 kilobases in length (13). With an estimated total of 10^{32} individual phage particles in the biosphere, they are considered the most abundant and genetically diverse biological entity on the planet (55). Found in all environments naturally inhabited by bacteria, phage numbers typically exceed 10^4 per millilitre of seawater and 10^9 per gram of soil in forests and agricultural areas (56). The therapeutic potential of phages was initially realised by Felix d'Herelle, where he demonstrated the use of intravenous phage treatment of invasive bacterial infections, and published his observations in 1931 (57). However, the lack of scientific rigor in early trials and the new found convenience of antibiotic administration saw interest in phage therapy in Western medicine diminish (3). Reignited due to the emerging antimicrobial resistance crisis, interest in phage therapy as an alternative to antibiotics has triggered a recent re-evaluation of its associated benefits and drawbacks (13). Understanding the structural characteristics of phages and their life cycles is crucial if the full potential of phage therapy is to be realised.

1.3.1 Phage morphology and life cycles

Introduced by Brenner and Horn in 1959, the method of negative staining for the visualisation of viruses through electron microscopy aids in classification of phages based on their morphology (58). There are tailed, polyhedral, filamentous and pleomorphic phages containing double stranded (ds) or single stranded (ss) DNA or RNA (59). Tailed, dsDNA phages belong to the order *Caudovirales* and are the most abundant and frequently isolated (59). Within *Caudovirales* exists 14 families constituting 96% of known phages: *Ackermannviridae, Autographiviridae, Chaseviridae, Demerecviridae, Drexlerviridae, Guelinviridae, Herelleviridae, Myoviridae, Podoviridae, Rountreeviridae, Salasmaviridae, Schitoviridae, Siphoviridae* and *Zobellviridae* (58, 60, 61) (Figure 1.1). The virions of the *Ackermannviridae, Herelleviridae* and *Myoviridae* families have contractile tails, whilst the remaining groups have non-contractile tails of varying length. None of their viral particles are enveloped (58). Outside of the *Caudovirales* order, there are several other families comprising polyhedral, filamentous and pleomorphic phages (58). These phages vary in their genome type and structural properties. In contrast to the *Caudovirales,* viral particles belonging to these groups lack tails and may

contain lipid or lipoprotein elements and may have lipoprotein envelopes (58). Tailed phages have shown the most promise in therapeutic application as they are the most understood (3).

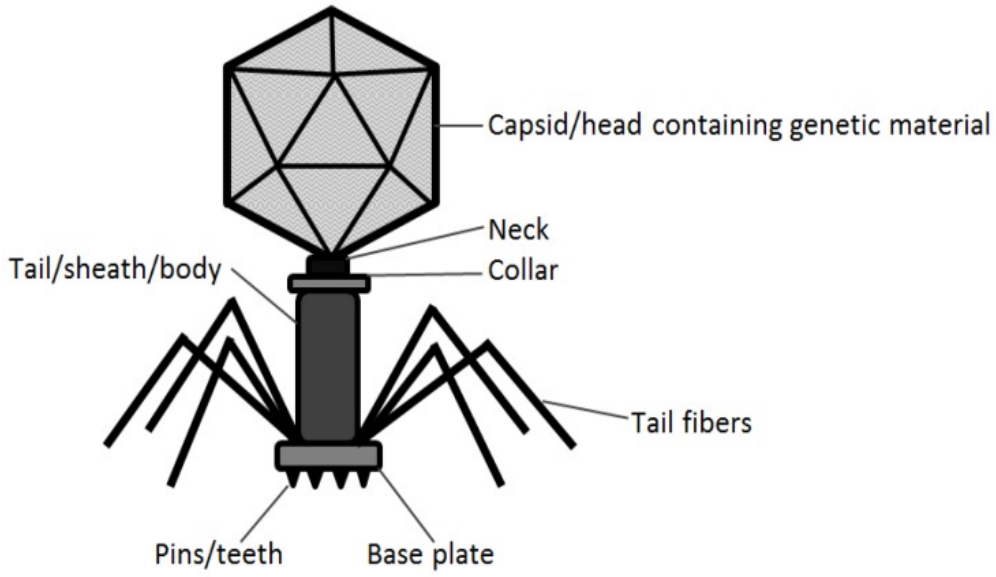

Figure 1.1 Schematic diagram of a bacteriophage belonging to the order *Caudovirales.* Illustrates generic structure of the capsid, containing genetic material and the base plate, harbouring receptor binding domains for attachment to the bacterial host. Doss *et al.* (12).

The relationship bacteriophages share with their host can be lytic or lysogenic (Figure 1.2) (13). In either case, the replicative machinery of the bacteria is hijacked to replicate the viral genome and translate viral proteins (62). The virulent, lytic relationship results in insertion of the viral genome into the host cell, immediate replication of the viral genome and subsequent construction of viral proteins (62). After the construction of new viral particles, lysis of the host cell is mediated by viral pore forming enzymes, a process which is initiated closely after infection. Host lysis results in release of viral progeny which then continue the cycle by infecting new host cells (62). However, in addition to the lytic relationship, temperate phages can undergo the lysogenic relationship, wherein the viral genome is integrated into the bacterial genome where it survives as a prophage (13).

Prophage propagation can occur via integration into the bacterial chromosomal DNA (using integrase or transposase enzymes) or as a self-replicating plasmid-like entity in the lysogen (lysogen: a bacterium containing a prophage) (13). The phage genome can then propagate within the lysogen until environmental stress triggers the lytic cycle. Temperate phages are not

ideal in phage therapy applications, as the prophage may not induce rapid lysis and can encode several protective factors to promote lysogen survival and prevent viral superinfection (63).

The specificity of a phage host range is an important attribute in its characterisation and is crucial to its therapeutic implementation (13). The tails of *Caudovirales* phages end with a base plate (Figure 1.1) containing receptor binding domains (RBDs) (64). RBDs are specific to cell surface receptors on bacterial hosts and dictate which bacteria the phage can attach to and infect (64). As RBDs are specific, the host range of lytic phages can be extremely narrow, allowing infection within a single bacterial strain only. Narrow host range coupled with the difficulty of isolating novel phages presents challenges in therapeutic application (13). As isolating phages specific to a target bacterial strain can be time-consuming, finding one in time to treat a bacterial infection might not be possible (65). This issue is partly overcome by the use of a 'phage cocktail,' or a combination of phages intended to cover a broader host range (65). An advantage of phage host specificity is the reduced risk of collateral damage stemming from the unintentional targeting of the host microbiota (13). A defining characteristic of phages, that is influenced by their morphology and lifestyle, is the enormous amount of horizontal gene transfer their genomes are subjected to.

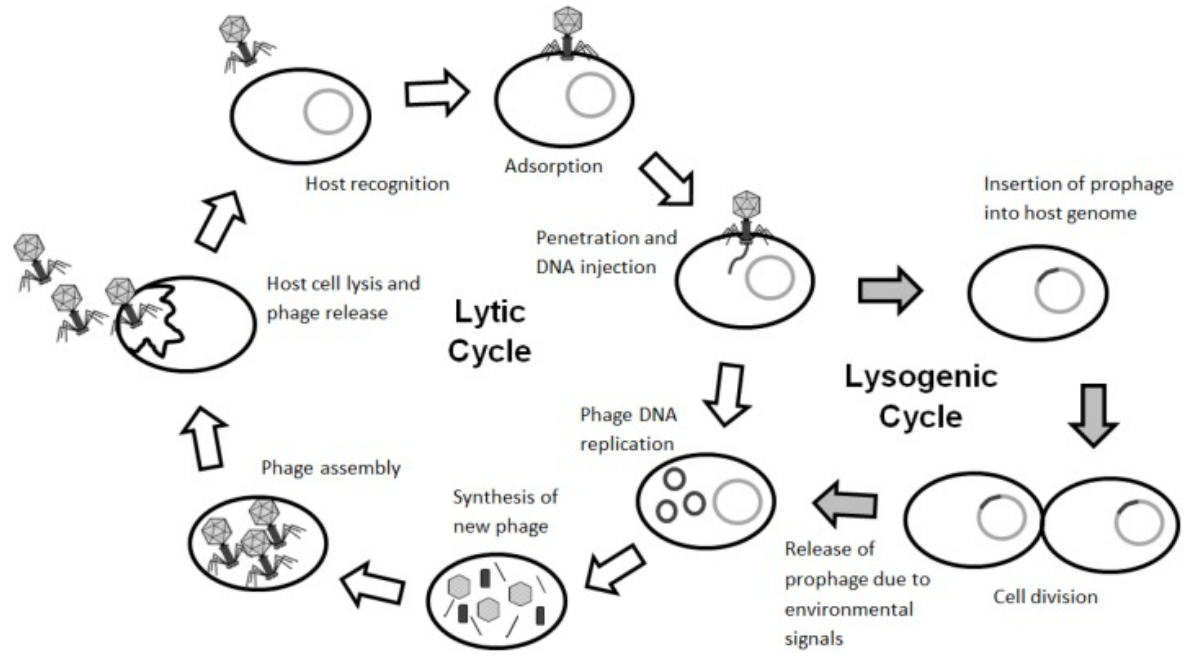

Figure 1.2 The lytic and lysogenic lifecycles of a phage. Doss *et al.* (12).

1.3.2 Horizontal gene transfer and phage genomes

The structure of phage genomes reflects the complexity of their evolutionary histories (66). Horizontal gene transfer (HGT) is an umbrella term for a number of processes that mediate sharing of genetic material between mature, independent cells (67). It is a driver of genetic diversity in prokaryotes and can occur within or across species. There are a number of mechanisms by which HGT can occur, including; transfer between two bacteria temporarily in contact (conjugation), transfer of a naked DNA fragment (transformation), and transport of bacterial DNA by bacteriophages (transduction) (67). HGT by transduction occurs constantly alongside lytic and lysogenic phage infection and heavily influences the evolution of both phages and their hosts (66). Recombination must occur for transduced genomic material to be preserved. In the context of phages, recombination is most often carried out by homologous recombination and nonhomologous (illegitimate) recombination. Nonhomologous recombination appears to happen randomly across phage genomes and genetic selection preserves only those recombinant genomes that produce viable phage progeny (68). Those resulting viable genomes are formed by modules, the boundaries of which demarcate functional genes. This modular exchange of genes results in the mosaicism observed in phage genomes. Whole genome/proteome approaches are proving to be vital to phage classification as marker gene phylogenies do not account for horizontal gene exchange of this nature (69).

1.3.3 Classification of phages

A number of unique challenges are faced when classifying novel phages (68). Classification is governed by the International Committee for Taxonomy of Viruses (ICTV) and, until recently, relied on morphological, nucleic acid type, and lifestyle characterisation (58). It is apparent that these approaches alone are not robust or efficient enough to handle the taxonomy of phages (68). The pervasively mosaic nature of phage genomes means that Linnaean, hierarchal phylogeny is ill-equipped to encapsulate their evolutionary history (68). Moreover, morphological characterisation is not always possible as numerous phage sequences are being deposited via metagenomic projects, or via extraction of prophages from bacterial chromosomes (70). As a result, a great deal of effort has been applied in developing an effective and scalable method of classifying phages. A 2020 report from the Bacterial and Archaeal Viruses Subcommittee of the ICTV addressed the challenges facing viral taxonomy by outlining an “ensemble of methods” suitable for the classification of novel phages (69). They

advised that classification should employ at least one established phylogenetic method and at least one whole genome-based method, and that additional methods may be required to strengthen discriminatory power. A combination of multi-marker gene phylogeny and network based-whole genome analysis balances statistically sound phylogenetic inference with contemporary and powerful visualisation. These contemporary approaches help to overcome the longstanding difficulties resulting from the complexity incorporated by HGT into phage genomes.

1.3.4 Network phylogeny

A review published by Hans-W Ackermann in 1998 reveals a sentiment about the factors that challenge protein coding sequence comparison as a technique for uncovering phage relatedness (71). Some of those factors being; the antiquity of phages means that any similarity detectable by amino-acid sequence alignments has been erased, the relatedness of two phages is likely undetectable without an intermediate phage capable of linking the two and, whole-genome amino acid comparisons cannot detect small regions of homology shared by distantly related phages. These are exactly the problems that gene sharing network-based approaches have begun to address. Advances in computing technology mean that huge datasets containing protein coding sequences from thousands of phages can be compared simultaneously. This allows for the generation of protein families, or groups of related protein coding sequences (72). When pairwise phage relatedness is ranked by the number of shared protein families among them, relationships between distantly related groups of phages become clearer. When approached in this fashion, a reticulated network of phage phylogeny appears that accounts for the influence HGT has on each phage's evolutionary history (72). This approach also compensates for the enormous amount of nucleotide sequence divergence that occurs in the genomes of rapidly evolving phages; a factor that can result in nucleotide sequence similarity being totally undetectable (69).

1.3.5 Aspects of phage application

The use of phages in clinical settings has garnered significant interest as the burden imposed by MDR pathogens on health care continually increases. Several studies have demonstrated the ability of single phages at reducing or eliminating bacterial pathogens from clinical and veterinary sources (Table 1). More common is the use of multiple phages mixed into a cocktail,

as this may have advantages over single phage treatment, including; increasing the number of strains or species the cocktail lyses, increasing the evolutionary pressure on the host which may result in host fitness trade-offs or re-sensitisation to antibiotics, and managing resistant bacterial outgrowth by targeting phage-resistant strains (73, 74). Examples of phage cocktails in action include a study by Yen *et al.,* who demonstrated using animal studies that polyphage prophylaxis is a promising approach for treatment of cholera caused by *Vibrio cholerae* (75). Forti *et al.* showed that their 6-phage cocktail was able to rescue mouse and wax moth models from MDR *Pseudomonas aeruginosa* respiratory and bacteraemia infections respectively (76). Nale *et al.* implemented an optimised 3-phage cocktail in controlling swine and poultry strains of *Salmonella* in *in vitro* and *in vivo* studies, supporting the potential usefulness of their phages in breaking the cycle of infection in agricultural settings. Further, bacterial density and virulence were shown to be suppressed by a 3-phage cocktail targeting *Salmonella typhimurium* on lettuce and cucumber, providing evidence that the phage cocktail holds promise as a natural disinfecting agent (77). Each study employed strategies to design and test the efficacy of their phage cocktails, with these strategies varying between studies (Table 1.)

Countering phage resistance is an important consideration in cocktail design; one strategy to account for this is a layered cocktail approach, also named the Step-By-Step method by Gu *et al.* (78). The Step-By-Step method involves isolating phages against phage resistant isolates as they arise and including them in the cocktail. This ensures that phages lytic to known resistant variants of the target bacteria are included in the original cocktail and pre-empts the development of phage resistance when the cocktail is applied clinically. Alternatively, selecting phages with diverse entry mechanisms may also result in a more robust cocktail.

Phages that are genetically or morphologically diverse from each other may have alternative methods of entry or replication. Selecting phages with diverse entry pathways can increase the chance of infection in the case of surface receptor mutation and reduce the likelihood of multi-phage resistant isolates emerging (73). The added evolutionary pressure resulting from this multi-faceted offence may also attenuate bacterial virulence (78), supporting the host immunity in clearing infection (79). In the absence of genetic characterisation, some studies incorporate diversity by including phages that are morphologically distinct from each other. This may suffice for *in vitro* studies; however clinical trials necessitate genetic characterisation of the phages.

Not only does whole genome sequencing (WGS) assist in the classification of phages for viral taxonomy, but it also allows for the identification of genes that may be harmful to the host organism. “Undesirable genes” are those within a phage genome that are involved in lysogeny, toxin production or antibiotic resistance (80). It is crucial that a phage therapeutic agent is incapable of lysogeny, as this minimises the risk of the dissemination of antibiotic resistance genes (ARGs) by HGT and maximises the chance of host lysis. Genetic characterisation in some instances serves as a precursor to Bacteriophage Recombineering of Electroporated DNA (BRED), that allows lytic derivatives of temperate phages to be engineered via precise removal of the repressor gene responsible for lysogeny initiation (81). The resulting phages may then be studied *in vitro* to measure their suitability for therapy via several “efficacy metrics” (Table 1).

Requisite to clinical application is the thorough characterisation of phage behaviour *in vitro.* Factors reported when measuring the effectiveness of a phage cocktail include host range, frequency of phage resistant mutants, time taken for resistance to emerge, phage-phage and phage-antibiotic synergism, *in vitro* efficacies (liquid infection assays and biofilm disruption), host virulence trade-offs and development of cross-resistance (77, 78, 82). Liquid infection assays are most often performed using OD_{600nm} measurements as a surrogate for bacterial growth in liquid culture in the presence of phage. The resulting bacterial growth curves can represent the phages capacity to suppress bacterial growth and indicate the time taken for phage resistance to emerge. Reports performing liquid infection assays to assess cocktail effectiveness are becoming increasingly common (Table 1), however they are yet to be published regarding phages infecting *S. marcescens.*

1.3.6 *Serratia marcescens* bacteriophages

Phages for *S. marcescens* thought to be suitable for therapeutics have been documented in the literature (83). Isolated from environmental water in Japan, three phages were characterised and displayed varying lifecycles and host ranges. Matsushita *et al.* (83), tested these phages against 23 nosocomial strains of *S. marcescens* (83). Measured via spot assays, two of the three phages were found to be effective at lysing 35% and 56% of the strains and were both strictly virulent. Evaluation of phage therapy suitability was conducted by measuring phage absorption rate, latency period and burst size, however liquid infection assays were not performed and the development of phage resistance in the host was not investigated. The two phages, labelled

KSP90 and KSP100 underwent DNA sequencing of their major capsid protein (MCP) genes. Phylogenetic analysis of the pre-MCP proteins against pre-MCP protein sequences of 39 related phages revealed relatedness to two phages of therapeutic candidature (83). One of the related phages, phiEco32, has demonstrated potential in treating acute mastitis in cows (84). The combination of having a broad host range and virulent life cycles presents these phages as potential therapeutic agents.

S. marcescens phages have the potential to be used therapeutically; this potential can be bolstered by combining them with existing therapies (82). The isolation of *S. marcescens* phage vB_SmaP_SALSA (SALSA) from sewage preceded its morphological and phylogenetic characterisation. Classed as a T7-like phage, SALSA was able to lyse 5 out of the 20 clinical *S. marcescens* isolates used by the researchers, including SM01. Typical of clinical *S. marcescens* isolates, SM01 harbours the chromosomal *ampC* gene, making it intrinsically resistant to β-lactam antibiotics. Liquid infection assays were performed to measure the growth characteristics of SM01 under multiple treatments, including ampicillin/sulbactam (SAM) treatment and phage SALSA application. When challenged with phage SALSA alone, SM01 suffered initial growth inhibition before developing resistance to the phage. However, despite being intrinsically resistant to the drug, researchers demonstrated that the combination of phage and SAM resulted in complete bacterial eradication. This study represents the utility in augmenting existing treatments with phage to overcome resistance mechanisms.

1.5 Conclusion

Overuse of antibiotics doesn't just lead to the production of resistant microbes; it can have other negative health implications in humans and insects (47, 85). Microbial imbalance due to antibiotic exposure can lead to persistent infection in humans and is a prominent medical issue (85). In addition, the use of antibiotics in combatting bee larvae disease (foulbrood) has been shown to perturb the gut microbiota of adult honey bees (47). This leads to reduced fitness and increased susceptibility to bacterial infection. In an insect of such economic and ecological importance, the continued use of traditional antibiotic therapy appears to be blunt and irresponsible. *S. marcescens* is demonstrably virulent towards insects, and is a documented contributor to the multifaceted pressure driving honey bee decline (10). Whilst phage therapy protocols for the effective treatment of insect diseases are yet to be established, the value they

might have is easy to imagine. Coevolutionary dynamics dictate that in a system where *Serratia* and a *Serratia* infecting phage exist, phage resistance will eventually emerge (86). In this case the value may come from the virulence trade-off associated with the cost of evolution towards resistance (87). In any case, phage therapy appears primed to have an essential role in the battle against antibiotic-resistant pathogens.

MDR pathogens such as *S. marcescens* threaten immunocompromised patients and insects of agricultural importance alike. Implementation of phage therapy in humans is very much in its nascent stages, with clinical trials of intravenously administered phages being approved by the United States Food and Drug Administration in 2019 (88). The immediate rewards may come from experimentation using animal models and applying those results in agriculture and wildlife conservation. Development of these applications will no doubt be expediated by contemporary phage characterisation and phylogenetic approaches; such approaches will be necessary to elucidate relationships within the labyrinth of phage sequence data being generated globally.

Table 1.1 Summary of various studies measuring the effectiveness of phage cocktails (CT) at killing bacterial pathogens of clinical and veterinary origin. In each case, the CT design considerations are listed, as well as the ways in which CT efficacy were measured (efficacy metrics). Additionally, if mentioned in the study, the time taken for the host to develop phage resistance post CT application is noted. "Liquid infection assays" refers to OD_{600} measurements taken as a surrogate for bacterial growth in liquid culture in the presence of phage.

Study	Target bacteria	Phages in CT	MOI	CT design strategies	CT efficacy metrics	Resistant bacterial outgrowth
(89) 2018	*A. salmonicidia (*Fish)	2	0.01	Genetic assessment Genetic diversity	Liquid infection assays	No (after 80Hrs)
(78) 2012	*K. pneumoniae* (Clinical)	3	N/A	Layered cocktail approach Host range complementation	Liquid infection assays Host fitness and virulence trade-offs Murine bacteraemia model	Yes (after 26Hrs)
(81) 2019	*M. abscessus* (Clinical)	3	N/A 10^9 I.V.	Bacteriophage Recombineering of Electroporated DNA (BRED – removal of repressor gene) Host range complementation	Efficiency of plating (EOP) Clinical metrics (Cystic fibrosis patient treated)	Yes
(76) 2018	*P. aeruginosa* (Clinical)	6	2.5	Morphological diversity Gene content analysis Host range complementation Transduction assays Genetic diversity	Liquid infection assays EOP Biofilm disruption assay Mouse respiratory model (MOI 0.3) Larvae model (MOI 25)	Yes (after <24Hrs)
(90) 2020	*K. pneumoniae* (Clinical)	3	6	Layered cocktail approach Receptor site diversity Morphological diversity Phage absorption quantitation	Liquid infection assays EOP	No
(91) 2021	*Salmonella spp.* (Poultry & swine)	3	10	Morphological diversity Host range complementation	Liquid infection assays Larvae model (MOI 10)	Yes (3-10Hrs)
(77) 2018	*S. typhimurium* (Fresh produce)	3	1	Receptor site diversity Plaque morphology Host range complementarity Cross-resistance examination Gene content analysis Morphological diversity	Liquid infection assays Hourly CFU counts Host fitness and virulence trade-offs Frequency of phage resistant mutants Food application	Yes (after 8-10Hrs)
(75) 2017	*Vibrio cholerae* (Clinical)	3	1	Receptor diversity (LPS 01 Antigen, OM Porin OmpU, 3rd unknown)	Liquid infection assays *In vivo* mouse & rabbit models	Yes (after 24Hrs)

				Previous characterisation of phages	Genetic based predictions of virulence trade-offs	
(80) 2020	*P. aeruginosa* (Clinical)	5	N/A 10^9 PFU	Host range complementation Layered cocktail approach Genetic diversity (dsDNA and dsRNA phages)	Liquid infection assays EOP Biofilm disruption assay	Yes (after 5 days)
(92) 2020	*E. coli* (Poultry)	6	<1		Liquid infection assays Biofilm disruption assay	No (after 24Hrs)
(93) 2020	Infantile sepsis blood pathogens (6 genera, gram negative)	29	10^6 PFU each	Plaque size and morphology *No genetic characterisation*	Liquid infection assays Spot assay	Not mentioned
(94) 2020	*S. aureus* (Clinical)	2	0.1-10	Gene content analysis Morphological diversity 2^{nd} phage in cocktail selected by plating against cocktail resistant strain. Phage titre optimisation	Only spot assays of CT performed Liquid infection assays and EOP performed for individual phages	Not mentioned

1.6 Research aims

Aim 1: Phylogenetic characterisation of the *Serratia* phages.

In the last decade, several comparative genomic studies have been performed on phages infecting various host groups (95-98). The ability to define relationships among large sets of genomic data has become increasingly accessible alongside the development of sequencing and computing technologies (99). This type of analysis is proving to be integral when characterising phages; however, it has never been performed on those infecting the genus *Serratia*. Despite showing promise as potential therapeutic agents, the number of *Serratia* phages in the National Centre for Biotechnology Information's (NCBI's) database is small relative to some host groups. An addition of six novel *S. marcescens* phage sequences (named Smarc1-6) is therefore substantial and warrants an effort to consolidate them into genetically related lineages. Thus, the main components of the first aim were to:

(i) generate a phage network phylogeny incorporating the six novel *S. marcescens* phages from our lab; and
(ii) generate clusters of related *Serratia* phages based on whole genome nucleotide similarity analysis and marker gene phylogeny.

Multiple phage characterisation approaches are utilised in this report. The purpose of these approaches is to obtain predictive, informative, and explanatory classifications that incorporate the phages true history of evolution.

Aim 2: Assess the therapeutic suitability of four novel phages (Smarc1-4) infecting *S. marcescens* strains of clinical origin.

Having been genetically characterised previously by members of my lab, the therapeutic suitability of four novel phages lytic against a panel of clinical *S. marcescens* isolates requires investigation *in vitro*. The main components of this second aim were:

(i) Monitor and compare the lytic ability of the phages against the *S. marcescens* strains in liquid culture;
(ii) Assess the lytic ability of cocktails containing different phage combinations; and
(iii) Investigate the phage dynamics when applied in the cocktails via qPCR.

Chapter 2: Materials and Methods

Please note: Sections 2.2.1, 2.3.2, 2.3.8 & 2.3.9 describe methods required to isolate phages and extract, manipulate and sequence phage DNA. These procedures were completed previously by members of the Petrovski Lab and were included in this document for the sake of completeness and to guide future work.

These methods describe the isolation and sequencing of 6 novel phages infecting the common host *S. marcescens.* The phages are called Smarc 1 – 6. This is followed by a description of the bioinformatic procedures involved in performing a comparative genomic analysis of these 6 Smarc phages against 26 *Serratia* phage genomes from the NCBI's public database. The methods describing the characterisation of the growth dynamics of *S. marcescens* monoculture and phage plus host coculture were only performed on Smarc1-4. This is due to Smarc5's inability to lyse in liquid culture and Smarc6's unavailability at the time of testing.

2.1 General Laboratory Procedures

2.1.1 Media, buffers, solutions and sterilisation

Media used and their respective compositions are listed in Appendix 1. Milli-Q H_2O was used as a solvent in solution preparation. Media, buffer, solutions and equipment were sterilised by autoclaving at 100 kPa, 121 °C for 20 min. Solutions unable to be autoclaved were sterilised using a 0.22 μm pore filter (Sartorius).

2.1.2 Bacterial strains and bacteriophages

Bacterial strains and bacteriophages used in this study are listed in Table 2.1. *S. marcescens* strains were incubated at 37 °C overnight on NA or in NB. All cultures were grown aerobically, and liquid cultures were incubated with shaking. All bacteria and bacteriophages were stored at 4 °C and stocked at -80 °C.

Table 2.1 Bacterial Strains and bacteriophages used in this study.

Bacterial Species	Strain	Source
Serratia marcescens	Y21	Chan, M – Clinical isolate
Serratia marcescens	G10	Chan, M – Clinical isolate
Serratia marcescens	B61	Chan, M – Clinical isolate
Serratia marcescens	Y37	Chan, M – Clinical isolate
Serratia marcescens	Y21 2.1	Chan, M – Clinical isolate
Serratia marcescens	Y21 3	Chan, M – Clinical isolate
Serratia marcescens	Y21 4	Chan, M – Clinical isolate
Serratia marcescens	B61 1	Chan, M – Clinical isolate
Serratia marcescens	B61 9	Chan, M – Clinical isolate
Serratia marcescens	Y37 2.2	Chan, M – Clinical isolate
Serratia marcescens	Y37 2.3	Chan, M – Clinical isolate
Serratia marcescens	Y37 23	Chan, M – Clinical isolate
Serratia marcescens	AM923	Marenda, M – Insect isolate
Serratia marcescens	17-177	Marenda, M – Insect isolate
Serratia marcescens	16-1099	Marenda, M – Insect isolate
Serratia marcescens	17-042	Marenda, M – Insect isolate
Serratia marcescens	17-065	Marenda, M – Insect isolate
Serratia marcescens	17-249	Marenda, M – Insect isolate
Serratia marcescens	18-593	Marenda, M – Insect isolate
Serratia marcescens	19QA	Marenda, M – Insect isolate
Serratia marcescens	19-015	Marenda, M – Insect isolate
Serratia marcescens	LH33P3	Marenda, M – Insect isolate
Serratia marcescens	BI12	Marenda, M – Insect isolate
Serratia marcescens	CENT 2.13	Marenda, M – Insect isolate
Bacteriophages		
Smarc1	-	Petrovski Lab – Environmental isolate
Smarc2	-	Petrovski Lab – Environmental isolate
Smarc3	-	Petrovski Lab – Environmental isolate
Smarc4	-	Petrovski Lab – Environmental isolate
Smarc5	-	Petrovski Lab – Environmental isolate
Smarc6	-	Petrovski Lab – Environmental isolate

2.2 Microbiology Procedures

2.2.1 Bacteriophage screening and enrichment

Environmental samples were collected from local sources. Liquid waste from worm farms were centrifuged at 6,000 *x g* for 10 min. The supernatants were filtered through a 0.22 μm pore filter and 100 μl was added to 10 ml log-phase broth of *S. marcescens.* The culture was then incubated as per the host's requirements overnight. Following incubation, the enrichment broth was centrifuged at 12, 000 *x g* for 10 min and the supernatant was filtered through a 0.22 μm pore filter.

To determine the titre of the phage filtrate, a serial dilution of the filtrate was prepared to 10^{-7} and spotted out on a host bacterium lawn plate to calculate the plaque forming unit (pfu). The pfu was calculated using the dilution factor where the number of countable plaques was between 30 – 50 with the following formula:

$$pfu = \frac{dilution\ factor\ \times\ number\ of\ plaques}{amount\ plated\ (10\mu l)}$$

If phages could not grow to high titre (> 10^6 pfu/ml) in broth, 100 μl of phage filtrate was spotted onto multiple host bacterium lawn plates and grown as per host requirements. The clearings were scraped using 1 ml of NB per plate and a sterile glass rod. The scrapings were centrifuged at 12, 000 *x g* for 10 min and the supernatant was filtered using a 0.22 μm pore filter. The titre was then determined by serial dilution and pfu calculation.

2.2.2 Smarc host range testing

Lawn plates of the bacterial strains were created using 2 colonies of each strain. Neat suspensions of each phage were dropped in 10 μl amounts onto the plates and incubated for 24 hr at 37º C. Strains that exhibited plaque formation/clearing were considered susceptible. Extent of clearing was visually assessed and graded.

2.2.3 Host strain CFU testing

In order to control the multiplicity of infection (MOI) in the phage liquid infection assays, the colony forming units/ml (CFU/ml) of the host strains were determined. For each strain, an

overnight broth culture was diluted to OD 0.05, and a serial dilution in H20 was performed to 10^{-6}. 100ul of dilutions 10^4, 10^5 and 10^6 were spread on separate NA plates in triplicate using a flamed rod and incubated overnight so that colonies grew in countable range (30-300 colonies). To determine CFU/ml the following formula was used:

$$\frac{CFU}{ml} = \frac{number\ of\ colonies}{dilution\ \times\ volume\ plated\ (100\mu l)}$$

2.2.4 Growth curve analysis

All *S. marcescens* monocultures and phage plus host cocultures were incubated at 37º C with shaking at 200RPM using the CLARIOstar (BMG Labtech) in 96-well flat-bottom cell culture plates (Greiner) with a final volume of 300 µl. Cultures were seeded at OD_{600} 0.05, measured using a Novaspec III+ Spectrophotometer (Biochrom), with coculture MOIs set at 0.1, 1 & 10. Subsequent OD_{600} measurements were recorded using the CLARIOstar plate reader platform, using the well-scan (22 flashes per well, 300 µl path correction) function, every 30 minutes for the duration of the experiment. Start point and endpoint well samples (20 µl aliquots) were immediately stored at -20º C until required for qPCR assay processing.

2.3 Molecular Procedures

2.3.1 Reagents, buffers and centrifugation

Reagents and buffers used for bacteriophage DNA extraction and gel electrophoresis are listed in Appendices 2 & 3. Small volume centrifugations were performed in a bench top centrifuge (Eppendorf), at 16,000 *x g* unless otherwise stated.

2.3.2 Bacteriophage DNA extraction

Pre-existing DNA was removed by the addition of 10 µl DNAse (1mg/ml), 10 µl RNAse (1mg/ml) and 10 µl $MgCl_2$ (500mM) to 1ml of phage filtrate before incubation at room temperature for 30 min. In order to precipitate the phages, 20 µl $ZnCl_2$ (2M) was added to the solution which was then vortexed and incubated in a 37º C water bath for 5 mins. Phages were pelleted by a 5 min centrifugation and the supernatant was discarded.

In order to degrade phage capsid protein and release DNA, the pellet was resuspended in a solution containing 8 µl EDTA (20mM), 10 µl SDS (0.5%), 1 µl Proteinase K (10 mg/ml), 140 µl NaCl (200mM) and 141 µl Milli-Q H_20 before being incubated in a 55º C water bath for 1 hr.

Phage DNA was precipitated out of solution by adding 1:1 volume (200 µl) phenol/chloroform/isoamyl alcohol to the solution before vortexing until completely opaque/white. The solution was then centrifuged at 13,680 x g for 3 mins and 150 µl of the aqueous top phase was transferred to a fresh microcentrifuge tube. 150 µl isopropanol was added and left at -20º C overnight.

The DNA was pelleted by a 10 min centrifugation at 12, 000 *x g a*nd the supernatant was discarded. The pellet was washed with 70% (v/v) ethanol and centrifuged again for 5 min at 12,000 x g. The supernatant was discarded, and the pellet was dried in a Savant DNA 120 Speedvac concentrator (Thermo Scientific) for 5 min and resuspended in 30 µl of Milli-Q water and stored at 4 °C.

2.3.4 DNA Quantification

DNA concentration was quantified with a Qubit 3.0 fluorometer (Thermo Fisher Scientific) as per manufacturer's instructions. Qubit reagent and Qubit buffer were combined to a ratio of 1:200 to make a working solution. 2 µl of DNA was added to 198 µl of Qubit working solution, vortexed and incubated at room temperature for 2 min. The sample concentration was then read and quantified by a Qubit 3.0 fluorometer (Thermo Fisher Scientific).

2.3.5 Polymerase chain reaction (PCR)

Concentration of reagents listed in Appendix 4.
Primers were designed to amplify specific DNA fragments in the phages genomes, using high fidelity polymerase chain reaction (PCR). The reaction mixture was prepared in a final volume of 50 µl and contained the following components: 5 ng template DNA, 0.5 µl iProof polymerase, 1.25 µl each of 20 mM forward and reverse primer, 1 µl 10 mM dNTPs, 10% total volume of DMSO, 5x buffer and Milli-Q water to make up 50 µl. The conditions of the high-

fidelity PCR are as followed in Table 2.2. The PCR fragments were subsequently purified using the Monarch® DNA Gel Extraction Kit (NEB).

2.3.6 Direct quantitative PCR (qPCR)

Primers were designed (Table 2.3) in Geneious Prime to amplify specific fragments of phage genomic DNA, using direct (DNA extraction not required) qPCR. Samples were thawed at room temperature and 20 µl reactions were prepared as per the QuantiNova™ SYBR® Green PCR Kit (Qiagen) manufacturer's instructions, using 1 µl of sample and forward and reverse primers at a final concentration of 250 nM each. The conditions of the qPCR experiments are detailed in Table 2.4. Quantification of qPCR product and subsequent melt curve analysis was performed using the CFX Connect Real Time PCR Detection System (BioRad). To prepare the qPCR standard, a 161 bp fragment of the 16s rRNA gene of *Pseudomonas aeruginosa* (AW60) was PCR amplified and purified. A dilution series of the purified 161 bp product (corresponding to a DNA copy number of $1.0 \times 10^2 – 1.0 \times 10^6$ µl $^{-1}$) were run in technical duplicates on all qPCR plates.

Table 2.2 High fidelity PCR experimental conditions.

Stage	Temperature	Duration	Cycles
Initial Denaturation	94º C	2 – 5 min	1
Denaturation	94º C	15- 30 sec	30 – 34
Primer annealing	54º C – 62º C	15 – 30 sec	30 -34
Primer extension	72º C	1 min 30 sec – 2 min	30 – 34
Final extension	72º C	5 – 7 min	1
Hold	4º C	-	-

Table 2.3 Primers used in this study.

Primer	Target Organism	Direction	Target Sequence	Product size (bp)	GC %	Tm °C
Phage Primers						
BP1	Smarc1	Forward	5′-GGCGAAAGTGACAAAGAC-3′	201	50	54
BP2		Reverse	5′-GTCATCGTATCCATTGGC-3′		50	53
BP3	Smarc2	Forward	5′-GCAGAACGATTTCCTGCA-3′	275	50	56
BP4		Reverse	5′-GAAAATCACCGATTCCGG-3′		50	54
BP5	Smarc3	Forward	5′-GGTGCGGAAGAAGTTCAA-3′	239	50	55
BP6		Reverse	5′-CCTCTTTGCGGCTAACAT-3′		50	55
BP7	Smarc4	Forward	5′-GTATGGTGGCGGTAAAAC-3′	234	50	53
BP8		Reverse	5′-GTTTTGCGGTTGCGTTGA-3′		50	58
Bacterial Primers						
1114F	*P. aeruginosa* AW60	Forward	5′-CGGCAACGAGCGCAACCC-3′	161	72	75
1275R		Reverse	5’-CCATTGTAGCACGTGTGTAGCC-3’		56	66

Table 2.4 Quantitative polymerase chain reaction experimental conditions.

Step	Time	Temperature	Ramp rate
PCR initial heat activation	2 minutes	95	Maximal/ fast mode
2- step cycling			
Denaturation	5 seconds	95	Maximal/ fast mode
Combined annealing/ extension	10 seconds	60	Maximal/ fast mode
Number of cycles	40	-	-
Melt curve analysis	-	-	-

2.3.6 Gel electrophoresis of DNA

DNA fragments of interest were visualised by agarose gel electrophoresis using a horizontal gel apparatus (Bio-Rad). Gels were made from 1% (w/v) agarose and 1 X TAE buffer and supplemented with SYBR safe DNA gel stain. Gels were submerged in 1 x TAE buffer and electrophoresed for 35 min at 100 V. To estimate the size of the DNA fragments, GeneRuler 1kb DNA ladder (Thermo Scientific) and λ DNA *Hind*III ladder (New England BioLabs) were used. To visualise and photograph the gels, the endure GDS gel documentation system (Labnet) was used.

2.3.7 PCR purification and recovery of DNA from agarose gel

The Monarch® DNA Gel Extraction Kit (NEB) was used for both PCR and gel purifications, as per manufactures instructions. For gel clean up, the DNA band of interest was excised from the gel using a sterile surgical scalpel under a UV transilluminator. The DNA band was placed in a sterile microcentrifuge tube with 400 µl of membrane binding solution per 10 mg of agarose. This mixture was incubated at 55 °C for 10 – 20 min until the band was fully dissolved. For PCR product purification, an equal volume of Membrane Binding Solution was added to the PCR product and mixed by pipetting.

The solution was then transferred to an SV Minicolumn attached to a Collection Tube and incubated for 1 min at room temperature. This was centrifuged at 16, 000 *x g* for 1 min. The flow-through was discarded and 700 µl of Membrane Wash Solution was added to the SV Minicolumn and centrifuged for 1 min at 16, 000 *x g*. The flow-through was discarded and the wash step was repeated but with 500 µl Membrane Wash Solution and centrifuged for 1 min at 16, 000 *x g*. The flow-through was discarded and the Minicolumn and Collection Tube were recentrifuged for 1 min at 16, 000 *x g*. The Minicolumn was then transferred to a sterile microcentrifuge tube and 30 µl of Milli-Q water was placed in the centre of the Minicolumn and incubated at room temperature for 1 min. The microcentrifuge tube with Minicolumn was centrifuged for 1 min at 16, 000 *x g* and the Minicolumn removed.

2.3.8 Next generation sequencing

DNA from bacteriophages of interest were extracted and sequenced using the Illumina MiSeq system. DNA sequences libraries were prepared using the QIAseq FX DNA Library Kits

(Qiagen) protocol as per manufacturer's instructions. DNA concentration was measured using a Qubit 3.0 Fluorometer (Section 2.3.4), and samples were diluted to 1 ng μl^{-1}.

2.3.9 Sequence assembly and annotation

Following sequencing, bacteriophage DNA sequences were subjected to in silico analysis to determine their genomics. CLC Genomics Workbench (version 9.5.4) was used to assemble the sequences and the Geneious Prime (version 2020.2.1) software was used to annotate.

2.4 Bioinformatic procedures

2.4.1 Phage Genome Sequences

NCBI's nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore) was queried for "serratia+phage" and the 26 verified *Serratia* phage genomes were downloaded individually. Only complete genomes were selected for analysis (the partial genome of phage Mtx was included as it was long enough to analyse). The 6 sequenced Smarc phage genomes from Petrovski Lab were collated alongside the 26 *Serratia* phage genomes in Geneious Prime to facilitate analysis. The accession number, description, source, location, year of isolation and genome size of each phage were recorded in Table 1.

2.4.2 Genome Clustering

Three genome comparison approaches were used to cluster related genomes. The 32 *Serratia* phage genomes were concatenated in order of smallest to largest and two identical FASTA files were generated using Geneious Prime. The two files served as the axes of a dot plot matrix generated using the Genome Pair Rapid Dotter program (GEPARD)(100) with parameters: word length=10; window size=0. This approach reveals patterns of similarity and dissimilarity for all genome pairs and allowed graphical identification of clusters of related sequences. Identical 10bp stretches between genomes present as diagonal lines, with the central traversing diagonal line appearing where each genome is identical to itself. Cluster criteria required that $\geq 50\%$ homology coverage is observed of the larger genome. That is, the cumulative pairwise nucleotide similarity must span $\geq 50\%$ of the larger genome at dot plot word length=10. Genomes that did not meet these requirements were classed as 'singletons.' Following this, the cluster designations were validated using the Virus Intergenomic Distance Calculator (VIRIDIC)(101). VIRIDIC calculates intergenomic distances using the BLASTN algorithm

with parameters: word size=7; reward=2; penalty=3; gap open=5; gap extend=2. Pairwise intergenomic similarities are displayed as percentage values and are coupled with genome alignment indicators. Finally, core protein phylogeny of 31 *Serratia* phage large terminase subunit (LTS) sequences was used to infer their evolutionary history. All phages in this analysis had been previously annotated. The LTS was identified as an appropriate phylogenetic marker gene as it was present in 31 of 32 genomes (an LTS sequence could not be identified for *Serratia* phage Knp4). The LTS sequences were extracted and aligned using the MUSCLE algorithm with default parameters (102). The Neighbour-Joining method was used to generate a consensus treed based on the Tamura-Nei genetic distance model and was visualised within the Geneious Prime software. The tree was rooted at *Serratia* phage Eta to facilitate comparison and branch support values were calculated from 100 bootstrap replicates. Meta data for each phage was overlayed onto the tree including cluster designation and year and location of isolation.

2.4.3 Genome maps

Phamerator (https://phamerator.org/) allows for pairwise genome alignments to be visualised using an online interface (103). To operate Phamerator, a database of the *Serratia* phage nucleotide sequences in GenBank format was created using the program PhamDB (104), which can either be run in Linux or as an image in Docker Desktop. The GenBank files were either downloaded from NCBI or generated using Geneious Prime. PhamDB stores BLASTN pairwise nucleotide comparisons and open reading frame (ORF) predictions based on BLASTP and CLUSTALW in a format accessible by Phamerator (103, 104). The resulting genome alignment maps were downloaded from Phamerator.

2.4.4 Network Analysis

A monopartite gene sharing network was generated using vConTACT2 (version 0.9.17) (105). Despite all of the *Serratia* phage genomes having already been annotated, the need for a single file containing all of the protein coding sequences in one format resulted in each genome being submitted to gene calling program MetaGeneMark (http://exon.gatech.edu/meta_gmhmmp.cgi) (106). The resulting amino acid FASTA files were combined. Within the CyVerse Discovery Environment, the gene-to-contig mapping program vConTACT2 Gene2Genome (version 1.1.0) assigned each protein coding sequence's coordinates to its contig ID for further analysis. The resulting file was coupled with the original

FASTA coding sequences file and submitted to vConTACT2 using the previously described optimal parameters (105). vConTACT2 incorporates NCBIs Bacterial and Archaeal Viral RefSeq (V85) genomes with ICTV approved taxonomy into the network as a reference database. In short, the Markov Cluster (MCL) algorithm consolidates every protein coding sequence from the input genomes and generates families of related proteins called protein clusters (PCs) (107). Subsequently, Diamond, an accelerated BLASTP-like program, determines the underlying protein-protein similarities within the PCs (108). Finally, ClusterONE generates viral clusters (VCs) of related genomes which are expressed as a network file and visualised in CytoScape (version 3.8.1) (109). The resulting network displays the *Serratia*, Smarc and NCBI reference phage genomes as nodes connected by weighted edges, where shorter edges represent phages that share more PCs. CytoScape allows for the creation of various 'styles' that illustrate different genome properties. In Fig. 3.4a., colours were assigned based on phage host preference to identify gene sharing between phage host groups. In Fig. 3.4b., taxonomic families were assigned colours to facilitate identification of interfamilial gene sharing and Fig. 3.4c. overlays cluster designations based on dot plot analysis.

2.4.5 Smarc Related Phages

In an attempt to taxonomize the novel Smarc phages, genomes within VCs inhabited by Smarc1-6 were identified in the gene sharing network. Each related phage genome was downloaded from NCBI and combined into a single file for VIRIDIC analysis. A heatmap of pairwise intergenomic nucleotide similarities and alignment indicators was generated and displayed in Fig. 3.3.

Section 3: Phylogenetics

3.1 Introduction

We have expanded the collection of phages infecting the genus *Serratia.* The National Centre for Biotechnology Information (NCBI) public database contains 25 verified *Serratia* phage genomes plus three partial genomes and ten unverified genomes (1st October 2020). Of the three partial genomes, one is long enough to be useful for analysis, bringing the total number of phage genomes used in this report to 26. Their genomes range from 38,678 bp to 357,154 bp in length with an average GC content of 45%. Prior to this study six novel *S. marcescens* phages were isolated and sequenced (Batinovic *et al* manuscript in preparation). Our addition of these six *Serratia* phages infecting the common host *S. marcescens* increases the total number of *Serratia* phage genomes that we can access to 32. Our *S. marcescens* phages, named Smarc 1-6, have genomes ranging from 40,056 bp to 60,348 bp with an average GC content of 56.9%. The 32 *Serratia* phages in this report (Table 3.1) belong to the order *Caudovirales,* with the most common family being *Myoviridae* (12), followed by *Siphoviridae* (11), *Ackermannviridae* (5), *Autographiviridae* (3) and *Podoviridae* (1). This section describes the establishment of clusters of related phage genomes based on various agreeable methods. Table 3.1 summarises information regarding the phages and includes accession number, description, isolation source (if published), country of origin, year isolated and genome size.

Table 3.1 Table of the 32 sequenced *Serratia* phages.

GenBank Accession No.	Description	Source	Location	Year	Genome size (bp)
Ackermannviridae					
NC_048736	*Ackermannviridae* **Vb_Smaa_3M**	River	Spain	1991	159,398
MF285619	*Ackermannviridae* **2050H1**	NA	Beijing	2017	159,631
Kx452697	*Ackermannviridae* **Knp4**	Sewage	India	2016	160,268
Nc_020083	*Ackermannviridae* **Phimam1**	NA	UK	2011	157,834
NC_042047	*Taipeivirus* **Vb_Sru_IME250**	*Serratia rubidaea*	China	2015	154,938
Autographiviridae					
NC_047774	*Teetrevirus* **Sm9-3y**	NA	China	2016	39,631
NC_047844	*Teetrevirus* **2050h2**	NA	China	2017	39,216
MN098329	*Teseptimavirus* **Pila**	Wastewater Treatment Plant	Texas	2018	38,678
Myoviridae					
NC_048668	*Myoviridae* **2050HW**	NA	China	2017	276,025
NC_048759	*Myoviridae* **Mtx***	Waste Water Treatment Plant	Texas	2017	68,621
NC_048792	*Myoviridae* **Moabite**	Mixed Swine Farm Samples	USA	2017	273,933
NC_048800	*Myoviridae* **Myosmar**	Pond Water	Texas	2017	68,745
MN334766	*Myoviridae* **PCH45**	Sewage	Dunedin NZ	2016	212,807
NC_041917	*Eneladusvirus* **BF**	Compost	Ireland	2017	357,154
NC_049464	*Tevenvirinae* **Muldoon**	Wastewater Treatment Plant	Texas	2020	167,547
Nc_024121	*Tevenvirinae* **Ps2**	NA	China	2014	167,266
MT176426	*Tequatrovirus* **Phizz30**	River	UK	2016	167,484
NC_041996	*Winklervirus* **CHI14**	NA	UK	2013	171,175

MF036691	*Winklervirus* **CBH8**	*Serratia sp.* ATCC 39006	UK	2014	171,175
Mf036692	*Winklervirus* **X20**	*Serratia Sp.* Atcc 39006	UK	2014	172,450
Podoviridae					
NC_048758	*Podoviridae* **Parlo**	Mixed Swine Faecal and Soil Samples	Texas	2017	62,853
Siphoviridae					
MN095770	*Siphoviridae* **Slocum**	Wastewater Treatment Plant	Texas	2018	112,436
MK608336	*Siphoviridae* **Serbin**	Pond Water	Texas	2015	42,882
MH553517	*Siphoviridae* **Scapp**	Wastewater treatment plant	Texas	2017	42,969
Nc_021563	*Siphoviridae* **Eta**	NA	Germany	1966	42,724
MN505213	*Chivirus* **Js26**	NA	NZ	2019	63,971
NA	**Smarc 1**	Worm farm effluent	VIC	2019	58,021
NA	**Smarc 2**	”	VIC	2019	44,871
NA	**Smarc 3**	”	VIC	2019	40,056
NA	**Smarc 4**	”	VIC	2019	43,034
NA	**Smarc 5**	”	VIC	2019	60,348
NA	**Smarc 6**	”	VIC	2019	59,531

* Partial genome.

3.2 Network Analysis

vConTACT2 network phylogeny produced a monopartite network of 2272 phage genomes, including six novel *S. marcescens* phages, 26 existing *Serratia* phages and 2240 NCBI reference phage genomes with ICTV approved taxonomy. The network includes phages infecting a variety of hosts and reveals gene sharing prejudice among host groups.

3.2.1 Gene sharing among host groups

Fig. 3.1. presents phage genomes as nodes, colour coded according to host genus. Host genera include *Serratia, Escherichia, Yersinia, Staphylococcus, Streptococcus, Lactococcus, Pseudomonas, Salmonella, Propionibacterium, Mycobacterium, Bacillus, Gordonia* and *Klebsiella* however this list is not exhaustive. Phages related at the protein coding level are

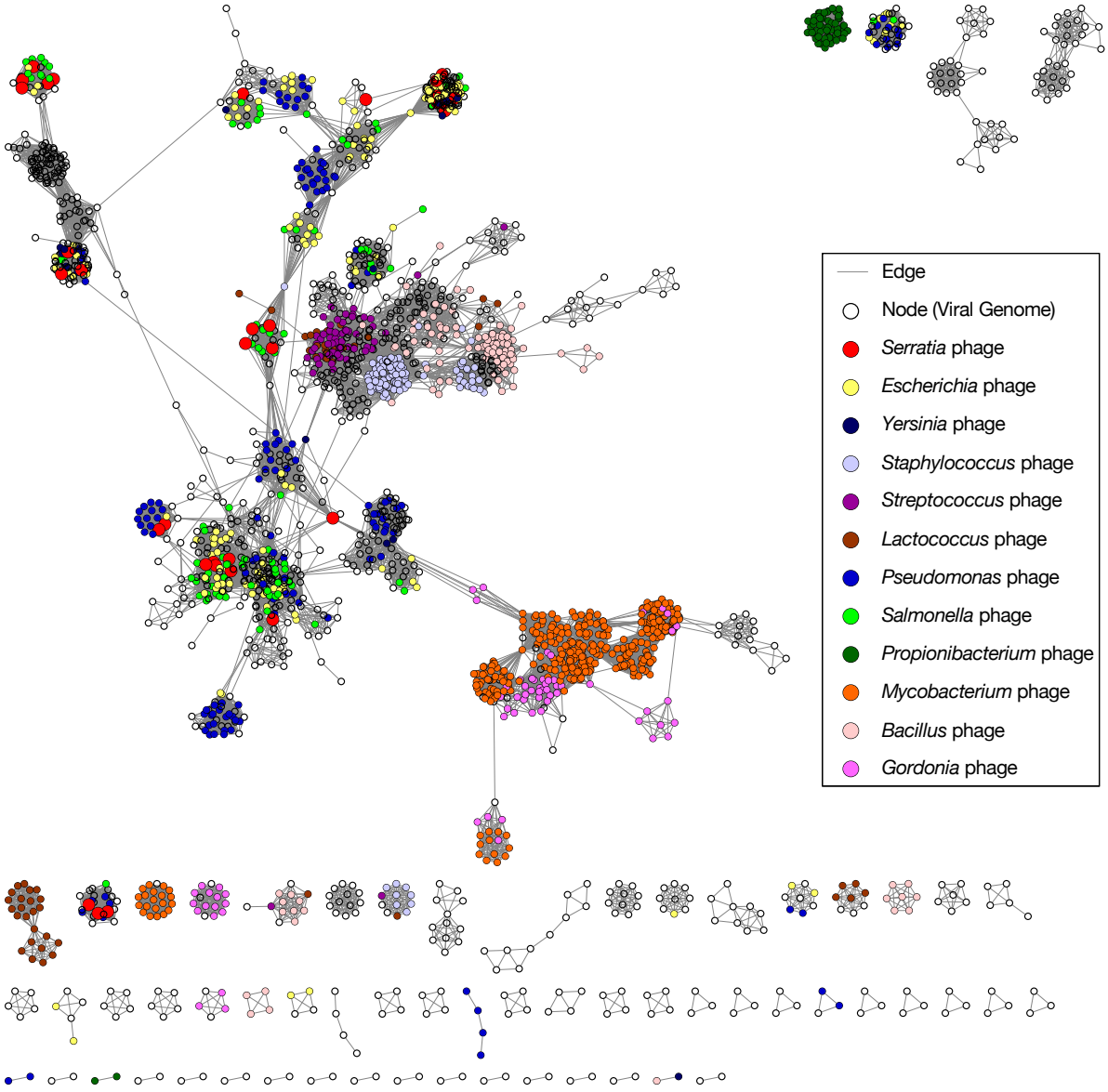

Figure 3.1 Network phylogeny of the *Serratia* phage genomes against a reference database. Nodes represent phage genomes with lines (edges) joining related genomes. Relatedness is determined as a function of the number of shared protein clusters (PCs) between genomes. Shorter edges indicate stronger relationships. Phages in NCBI's public phage database served as reference genomes. Only reference phages with ICTV approved taxonomy were included. Image shows nodes colour coded by the host genus they infect. The *Serratia* phages (red) are spread out among various VCs that include phages infecting multiple different hosts.

joined by edges, where shorter edges indicate greater relatedness. Groups of related phages form viral clusters (VCs). Isolated VCs tend to be dominated by a main host genus. Superclusters of highly connected phages form internal clusters segregated by host genus. This indicates that the gene sharing that leads to protein coding sequence similarity is influenced by phage host preference.

3.2.2 Gene sharing among taxonomic families

In addition to being influenced by host preference, network phylogeny indicates gene sharing is heavily restricted by morphology/taxonomic family. Fig. 3.2. shows that *Siphoviridae* phages dominate the central supercluster and are the most prevalent taxonomic family in the network. Taxonomic families rarely cross cluster demarcations, with VCs containing more than one taxon in only a few cases. There is no apparent correlation between the dominant taxonomic family of a VC and its level of genetic isolation, as *Siphoviridae* phages appear in the most and least interconnected regions of the network.

3.2.3 Genetic isolation among host groups

Varying levels of genetic isolation are observed between different host groups. The *Serratia* phages are dispersed across a number of interconnected VCs that contain predominantly Proteobacteria (*Salmonella, Escherichia* and *Pseudomonas*) phages (Figure 3.1). These VCs are distributed throughout the network and are highly interconnected, in contrast to the Actinobacteria (*Mycobacterium, Gordonia* and *Propionibacterium*) phages that are more contained and isolated from each other. Despite outnumbering the other phages, the *Mycobacterium* phages remain totally isolated except for a group of *Gordonia* phages that connect them to the Proteobacteria phage supercluster. An exclusively *Propionibacterium* phage cluster in the top right corner of the network represents the most genetically isolated host group. A second supercluster comprising of predominantly Firmicutes (*Streptococcus, Staphylococcus* and *Bacillus*) phage VCs exists and is connected by several edges to the Proteobacteria phage supercluster. As expected for Siphoviridae phages infecting a host belonging to the phylum Proteobacteria, the Smarcs group among VCs within the network's Proteobacteria supercluster.

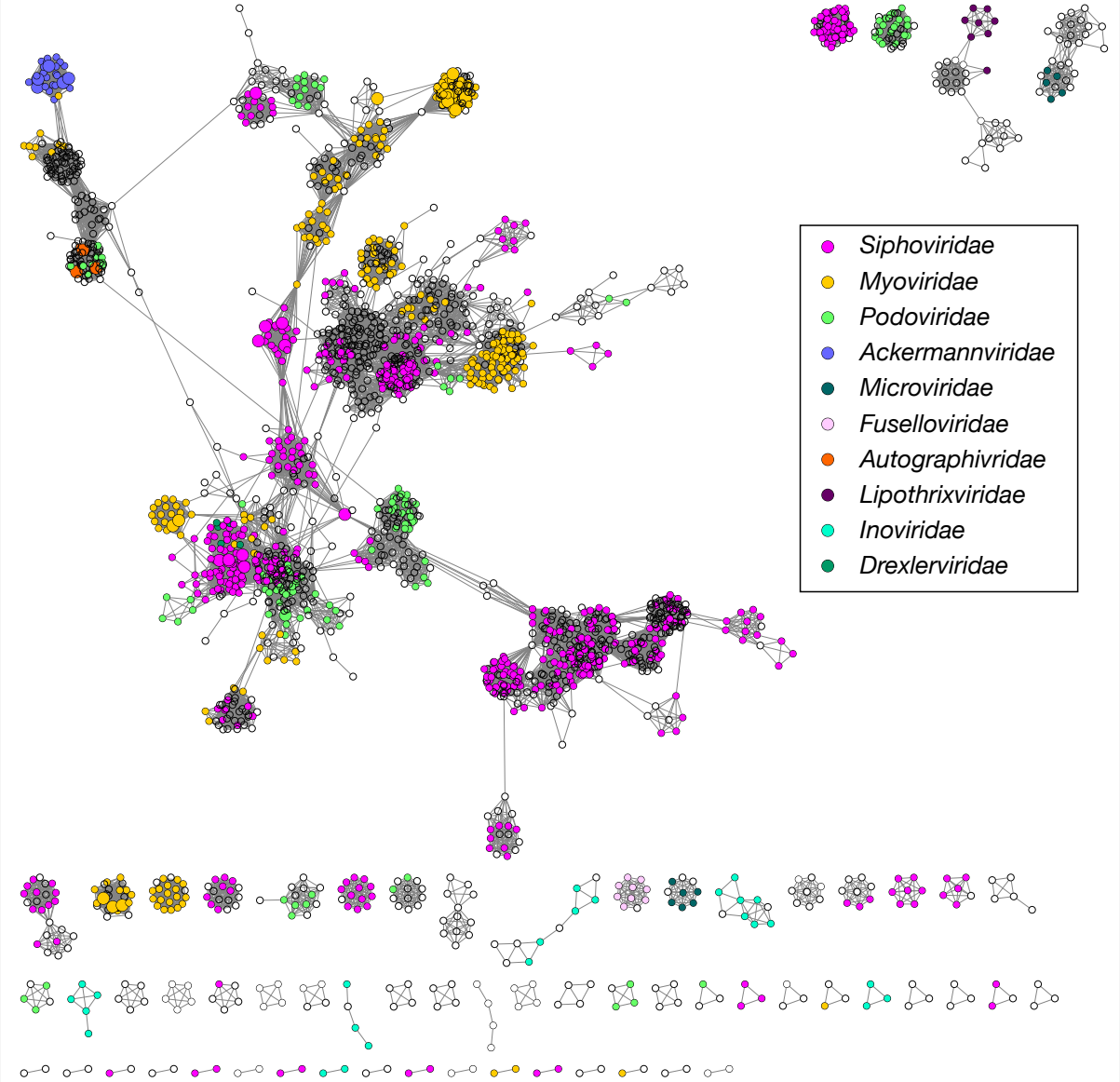

Figure 3.2 Network phylogeny of the *Serratia* phage genomes against a reference database. Nodes represent phage genomes with lines (edges) joining related genomes. Relatedness is determined as a function of the number of shared protein clusters (PCs) between genomes. Shorter edges indicate stronger relationships. Phages in NCBI's public phage database served as reference genomes. Only reference phages with ICTV approved taxonomy were included. Image represents phages coloured by their taxonomic family. *Siphoviridae* (purple) is the most prevalent and dominates the central supercluster.

3.2.4 Smarc related phages

The 46 phage genomes belonging to VCs inhabited by Smarcs 1-6 were grouped into a single nucleotide FASTA file for VIRIDIC analysis (Figure 3.3). Smarc1, 5 & 6 and *Serratia* phage JS26 belong to an exclusively *Siphoviridae* VC dominated by *Salmonella* phages of the genus

Chivirus. The Smarc1, 5 & 6 VC presents as a clearly defined, interrelated cluster and is only connected by single bridging phages to the central supercluster of *Siphoviridae* phages and to an outer cluster of *Myoviridae* phages. Intergenomic similarity calculation (Figure 3.3) showed that Smarc 6 shares 93.1% similarity at the nucleotide sequence level with *Salmonella* phage Chi of the genus *Chivirus.* Smarc5 shares 60.1% similarity with *Salmonella* phage BP12C of the same genus. Despite demonstrating protein coding sequence similarity, Smarc1 does not exceed 11% nucleotide sequence similarity with the other phages in its VC. Smarc2, 3 & 4 and Scapp & Eta exist centrally within the *Siphoviridae* supercluster which is well connected to smaller VCs of *Myoviridae* and *Podoviridae* phages. The majority of phages in this cluster are *Salmonella* and *Escherichia* siphophages belonging to the genera *Jerseyvirus* and *Dhillonvirus* respectively. Despite this, Smarc3 shares the highest nucleotide similarity with *Enterobacter* phage phiEap 2 at 38%, reflecting the heterogeneity of host genera in this supercluster. Smarc2 & 4 do not share significant similarity at the nucleotide sequence level with any of the phages in their VC.

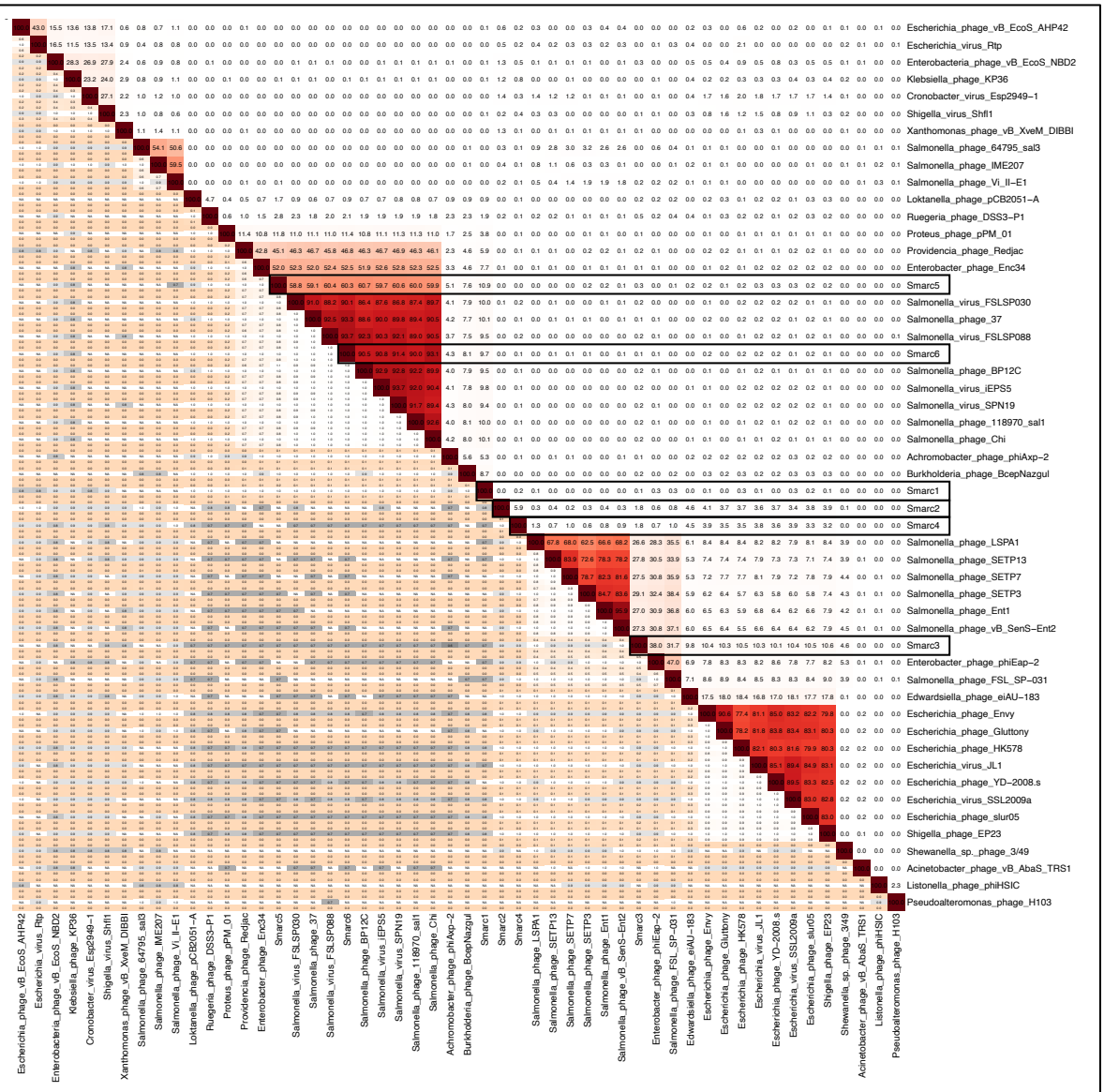

Figure 3.3 Intergenomic sequence similarity heatmap and alignment indicators of 52 Smarc related phage genomes. Percentage values of nucleotide sequence similarity calculated by VIRIDIC are displayed in top right half of graph. These values are overlayed with a heatmap where darker red indicates stronger sequence similarity. Each similarity value is complemented by three corresponding alignment indicators displayed in the bottom left of graph. The alignment indicators are as follows; Top & bottom represent the aligned genome fraction of the genome pair for the phage in this row and column respectively (darker colours represent lower values, or genome pairs that are less well aligned). The centre value represents the genome length ratio (darker colours indicate lower values, or genomes that are less similar in length). In this way, simultaneous visual assessment of a genome pairs alignment ratio, length ratio and percentage intergenomic similarity can be performed. Intergenomic sequence similarity of the 52 phage genomes that share VCs with the Smarcs in the gene sharing network (Figure 3.1). Smarcs 5&6 display 59 to 93% similarity to a group of *Salmonella* phages. Smarcs 1-4 do not share significant nucleotide sequence similarity with other phages from their VCs.

3.3 Clustering of *Serratia* Phage Genomes

3.3.1 Genome clustering

Intergenomic sequence similarity calculation was coupled with dot plot analysis to cluster related genomes. The 32 *Serratia* phage genomes were concatenated in order of shortest to longest and aligned along each axis of a dot plot matrix to facilitate clustering (Figure 3.4b). 20 phages were assigned to 8 clusters (Table 3.2) named A to H as follows: Cluster A (n=3), Cluster B (n=2), Cluster C (n=2), Cluster D (n=2), Cluster E (n=4), Cluster F (n=2), Cluster G (n=3), Cluster H (n=2). The remaining 12 phages did not meet clustering criteria and remain as singletons.

Table 3.2 *Serratia* phage cluster designations. Clusters range from A-H and are based on nucleotide intergenomic similarity analysis. 20 out of the 32 *Serratia* phages were assigned clusters while those that do not meet clustering criteria remain as singletons.

Clusters								
A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**Singletons**
Smp-3y	Scapp	Smarc5	Mtx	Vb_SmA-3M	Muldoon	CHI14	2050HW	PCH45
2050H2	Smarc4	Smarc6	Mysomar	2050H1	Ps2	CBH8	Moabite	BF
Pila				Phimam1		X20		Slocum
				Knp4				Serbin
								Parlo
								Eta
								Js26
								Smarc1
								Smarc2
								Smarc3
								PhiZZ30
								vB_Sru_IME250

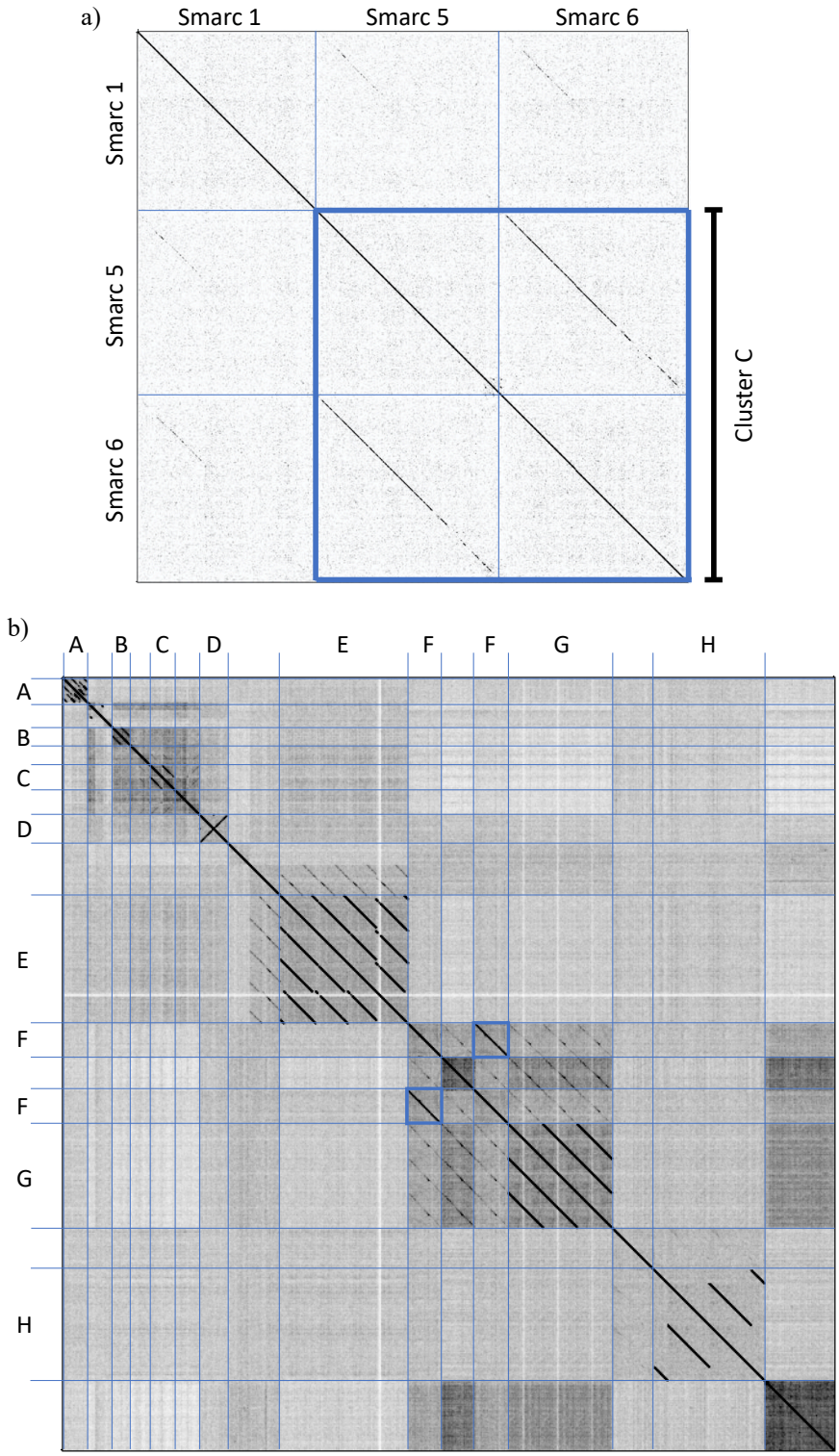

Figure 3.4 Dot plot analysis of 32 *Serratia* phage genomes. Dot plot analysis visualises sequence homology and facilitates clustering of related genomes. Analysis performed using Gepard (word length=10, window size=0). Areas of sequence homology are represented by a black line. Image **(a)** illustrates genome clustering criteria. Genomes are clustered if sequence homology spans >50% of the longer genome. Smarc5 & Smarc6 are obviously very similar and belong to cluster C. Smarc1 shows some homology to Smarc5&6 but does not meet the criteria and remains a singleton. Image **(b)** displays an all-against-all dot plot of 32 concatenated *Serratia* phage genomes in order of smallest to largest. Clusters of related genomes are visible and have been marked with blue lines and labelled.

This analysis was supported by VIRIDIC intergenomic sequence similarity calculation that revealed a minimum of 56% similarity between pairs of clustered genomes (Figure 3.5). The average pairwise intergenomic similarity among clustered genomes was 81%. In addition, alignment ratios and genome length ratios for all pairs of clustered phages returned nominal values. In support of these clusters being genetically related lineages, a phylogenetic tree was generated using an alignment of the 31 *Serratia* phage large terminase subunit sequences (Figure 3.6). The branch lengths within the cluster designations were short relative to those between clusters. These cluster designations were overlayed onto the gene sharing network (Figure 3.7) to illustrate the validity of the network.

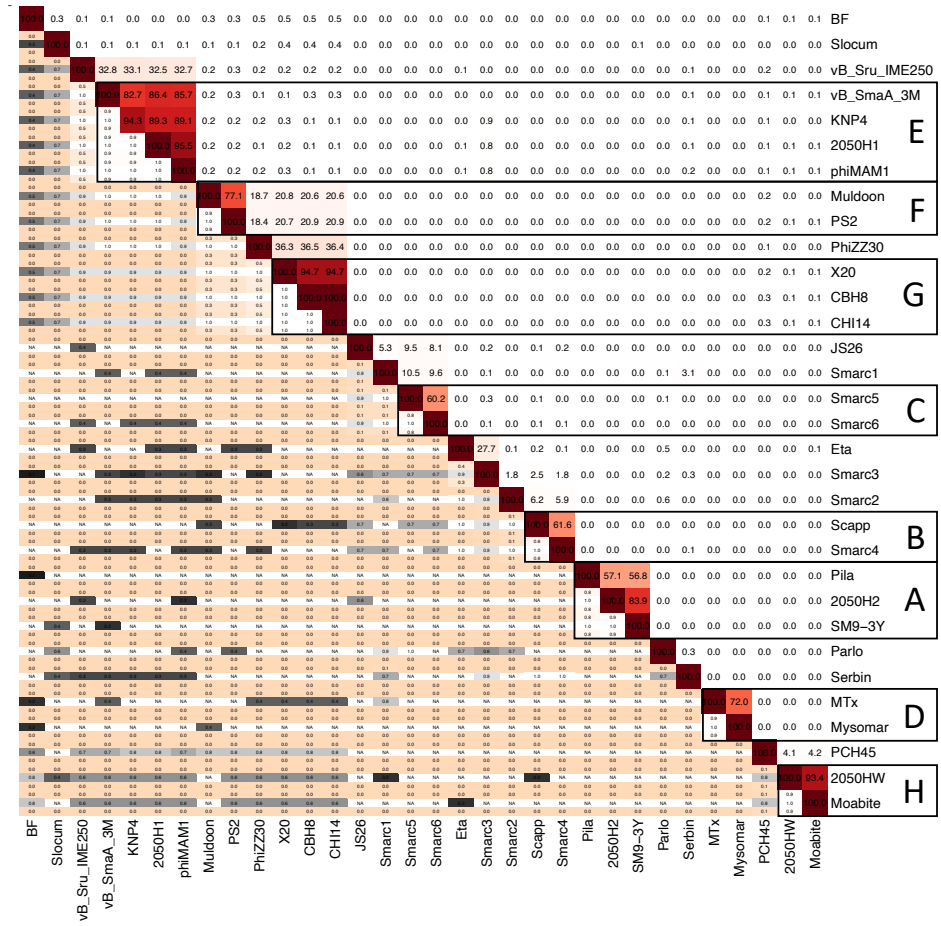

Figure 3.5 Intergenomic sequence similarity heatmap and alignment indicators of 32 *Serratia* phage genomes. Boxes denote cluster designations. Heatmap scale and alignment indicators explained in Fig 3.3

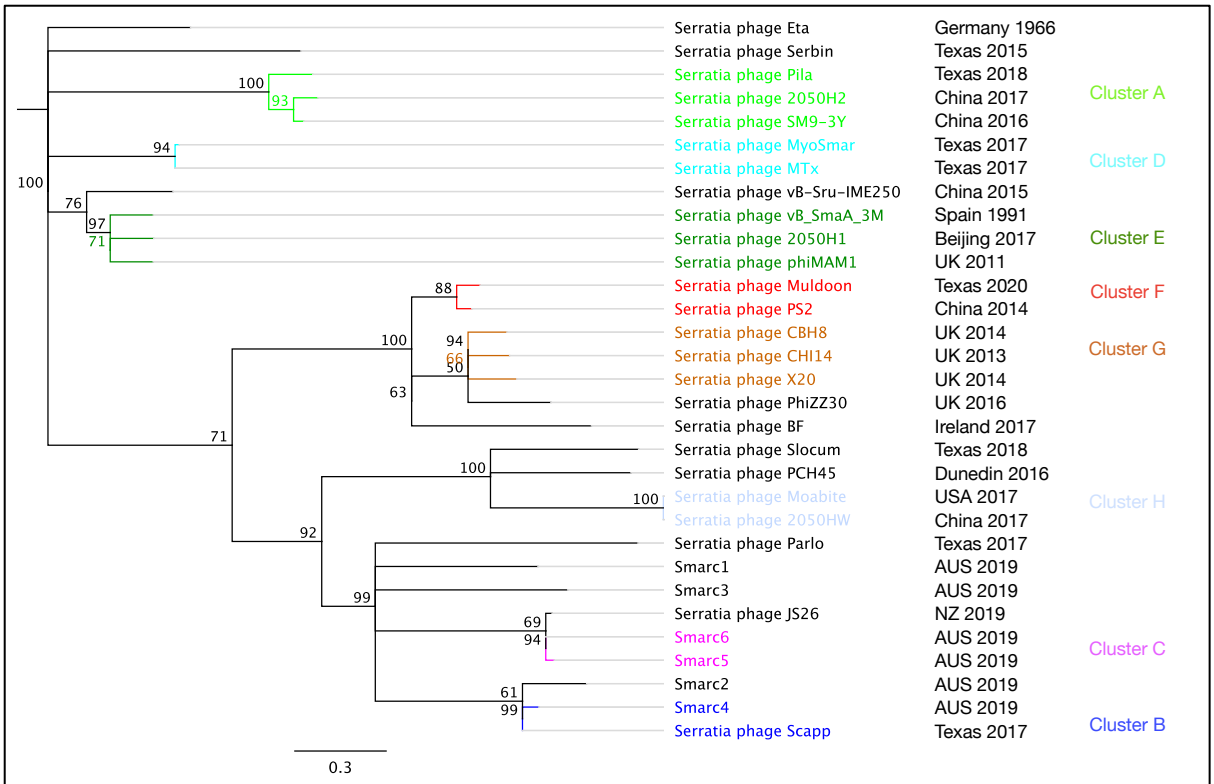

Figure 3.6 Marker gene phylogeny of 31 *Serratia* phage genomes. Neighbour joining tree based on MUSCLE alignment of large terminase subunit sequences. Labels coloured according to cluster designations based on dot plot. Branch lengths within cluster designations support that these are genetically related lineages. Year and location of isolation are listed to the right. Trees were rooted at *Serratia* phage Eta to facilitate comparison. The scale bar represents the number of nucleotide substitutions per site and branch support values were calculated from 100 bootstrap replicates. Note that phage KNP4 was omitted from the tree as a large terminase subunit sequence could not be identified.

3.3.2 Comparison of characterisation approaches

As network phylogeny is a relatively novel approach, agreement with alternative phage classification approaches is important. All cluster designations based on dot plot analysis agreed with VC groupings within the gene sharing network (Figure 3.7). Cluster F and Cluster G phages, despite not meeting clustering criteria, share significant enough protein coding sequence similarity to belong to the same VC. Likewise, singletons Smarc1, Smarc2, Smarc3, vB_Sru_IME250, Eta, PCH45, JS26 and PhiZZ30 grouped into VCs containing phages belonging to a cluster designation. The remaining singletons were dispersed across other Proteobacteria VCs within the network.

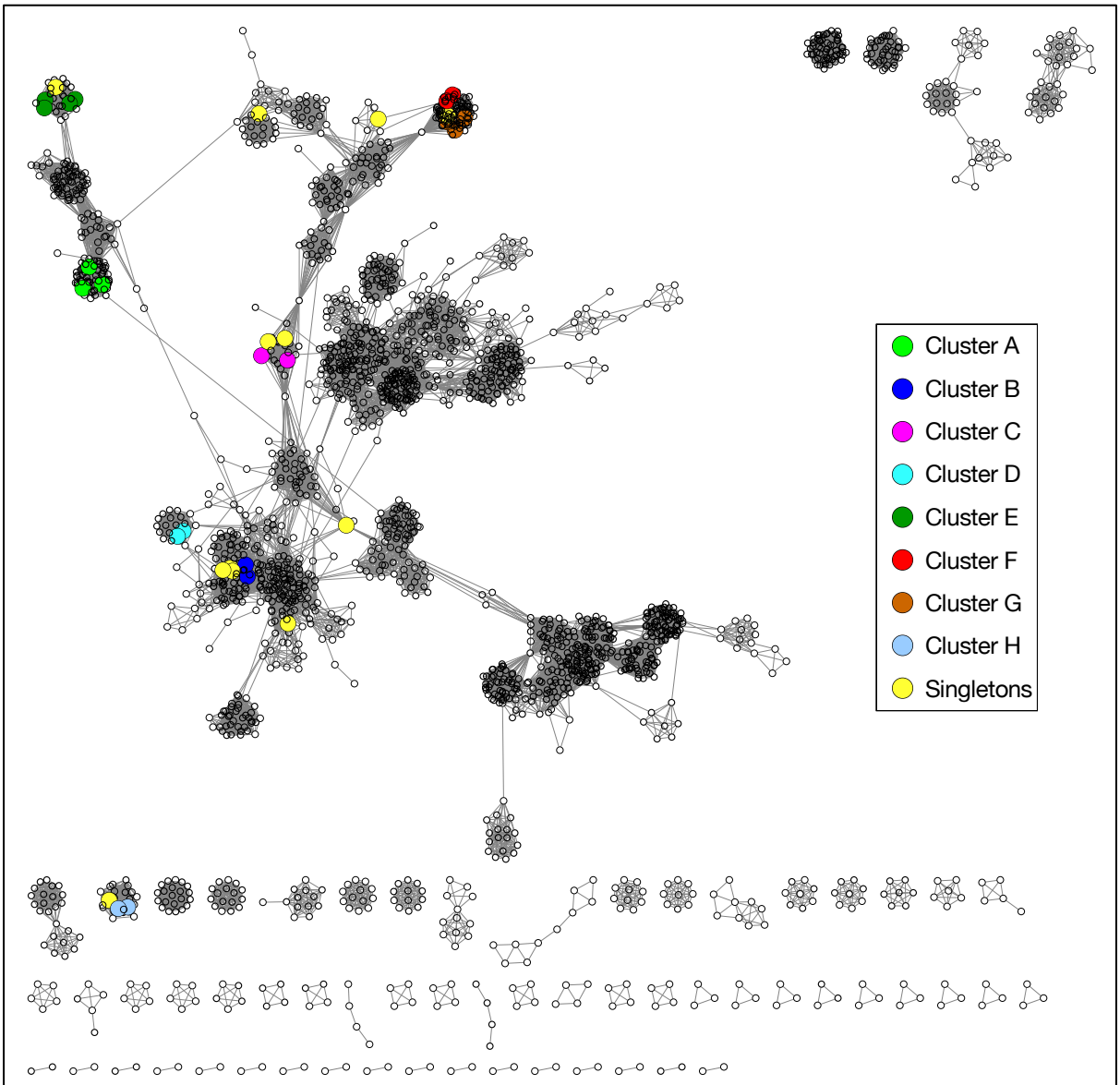

Figure 3.7 Network phylogeny of the *Serratia* phage genomes against a reference database. Nodes represent phage genomes with lines (edges) joining related genomes. Relatedness is determined as a function of the number of shared protein clusters (PCs) between genomes. Shorter edges indicate stronger relationships. Phages in NCBI's public phage database served as reference genomes. Only reference phages with ICTV approved taxonomy were included. Nodes representing the *Serratia* phages are coloured according to the cluster designations determined by nucleotide sequence similarity and marker gene analysis. The cluster designations correlate with the network VCs. 7 out of 12 singletons group into VCs that include *Serratia* phages that have cluster designations.

3.4 Genome maps

Inter-cluster pairwise genome alignments were performed and visualised using the Phamerator platform (Figure 3.8). Stretches of nucleotide sequence homology are represented by purple

shading and are interrupted by abrupt lengths of dissimilarity. Predicted ORFs are represented by boxes. Stretches of sequence homology appear to align with gene boundaries or at the boundaries of groups of genes. Genome synteny is conserved to varying degrees between pairs of phages. Cluster A phages, Pila and SM9-3Y, share two sections of extended homology with matching gene order that have been switched; a similar pattern is observed between Cluster D phages Mtx and Mysomar. The remaining genome pairs demonstrate conserved gene order with discrete interspersed regions of dissimilarity.

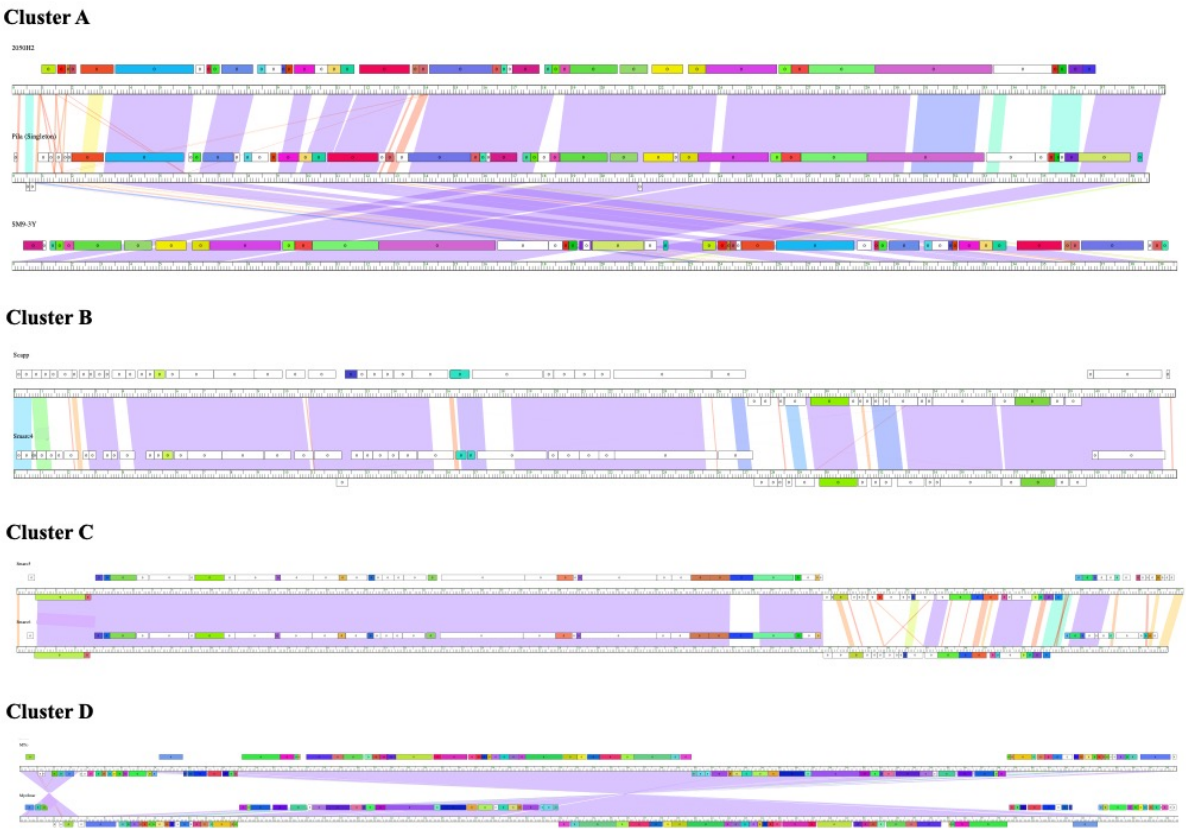

Figure 3.8 Pairwise alignment of clustered *Serratia* phage genomes. *Serratia* phage clusters A – D are displayed as Phamerator genome maps where shading indicates nucleotide sequence homology between pairs of genomes. Homology is calculated using BLASTN and colour coded according to the colour spectrum, where purple is the strongest. No shading indicates a lack of homology with an E value of 1e-4 or better. Boxes indicate predicted ORFs based on BLASTP and CLUSTALW, where related ORFs share the same colour. Boxes are either above or below, representing the transcription directions forward and backward respectively. Order of phages in each cluster is: (A) 2050H2, Pila & SM9-3Y; (B) Scapp & Smarc 4; (C) Smarc 5 & Smarc 6; (D) Mtx & Mysomar. The remaining clusters were not pictured as the length of the genomes precluded effective visualisation.

Section 4: *In vitro* Characterisation of Smarc phages

4.1 Introduction

Of the six *S. marcescens* phages isolated and sequenced by our lab, four were suitable for *in vitro* characterisation and assessment of their therapeutic suitability. Despite lysing their respective hosts on solid media, Smarc5 & 6 displayed no lytic activity in liquid culture and were omitted from further testing. Morphological and genetic characterisation of Smarc1-4 had been performed previously, and it can be reported that their genomes are lacking genes involved in toxin production, lysogeny and antibiotic resistance (S. Batinovic manuscript in preparation). This presents them as potential candidates for use in clinical settings and as biocontrol agents. As such, we have assessed their lysing ability against a representative panel of four clinical *S. marcescens* isolates. To address the issue of phage resistant bacteria, eight phage resistant mutants were generated from the panel of clinical *S. marcescens* isolates and included in this analysis (Table 4.1) (S. Batinovic manuscript in preparation). An additional 12 strains of insect origin (Supplied by Dr Mark Marenda, University of Melbourne) were included in the host range analysis to complement the host range testing.

This section describes the development of a liquid infection assay workflow that measures the growth of the clinical *S. marcescens* strains and their sensitivity to the four Smarc phages, individually and in cocktails. Optical density at 600nm (OD_{600nm}) measurements are used as a surrogate for bacterial growth and qPCR is used to quantify the change in abundance of each individual phage when they are applied as a cocktail.

4.2 Phage resistant *S. marcescens* strains

Phage resistant mutants were generated for *S. marcescens* strains Y21 (Smarc1 host), B61 (Smarc2 host) and Y37 (Smarc4 host). *S. marcescens* strain G10 was susceptible to Smarc2 and after numerous attempts no mutants were isolated. Each *S. marcescens* mutant was whole genome sequenced and mutation detected by alignment to the wildtype sequence.

The mutations identified by genome alignment exist within either hypothetical proteins of unknown function, transcriptional regulators or phosphomannomutase and phosphatase enzymes (Table 4.1). Two of the mutations, belonging to Y37 mutants 2.2 & 2.3, occurred upstream of promoter regions and did not have a discernible effect. None of the mutations identified here occur in genes responsible for phage receptor proteins, however, they may have consequences affecting receptor protein expression as they inhibit lysis of the host.

Table 4.1 Phage resistant mutants. For each Smarc host (except G10) phage resistant mutants were generated and sequenced to identify the type and location of each mutation.

Strain	Mutant	Mutation	Effect	Position	Protein product
Y21 (Smarc1 host)	**WT**	-	-		-
Y21	2.1	Deletion	Frameshift c.319delA p.Met107fs	1,672,962	Phosphomannomutase CpsG
	3	SNP	Missense variant c.998C>T p.Thr333Ile	1,655,863	Hypothetical protein
	4	Deletion	Frameshift c.65delT p.Leu22fs	1,654,923	Hypothetical protein
G10 (Smarc2 host)	**WT**	-	-		-
B61 (Smarc3 host)	**WT**	-	-		-
B61	1	Deletion	Frameshift c.94delA p.Ile32fs	1,590,429	Protein tyrosine phosphatase
	9	Deletion Insertion	Frameshifts c.94delA p.Ile32fs, c.156dupG p.Pro53fs	1,590,429 & 3,777,517	Protein tyrosine phosphatase & efflux periplasmic adaptor subunit
Y37 (Smarc4 host)	**WT**	-	-		-
Y37	2.2	SNP	NA	3,189,228	N/A
	2.3	SNP	NA	1,595,692	N/A
	23	SNP	Missense variant c.524A>T p.Lys175Met	3,189,848	Transcriptional regulator RcsA

4.3 Host range analysis

Key indicators of suitability for phage therapy are phage host range and lysing ability. To assess the lytic range of Smarc1-4 and to inform the liquid infection assay experimental design, a panel of 23 clinical and insect *S. marcescens* strains were subjected to phage spot assays. The phage resistant derivatives of Y21, B61 and Y37 were included to monitor the development of cross-resistance arising from mutation, however none was observed. Lysing ability was visually assessed as the clearing produced when a drop of neat phage suspension was placed on a bacterial lawn (Table 4.2). Smarc2 demonstrated the broadest lysing ability, forming plaques on 70% of the strains. Smarc4 lysed 61% of the strains, followed by Smarc1 at 35% and Smarc3 at 9%.

Table 4.2 Smarc host range data. Preliminary testing of Smarcs lysing ability was carried out against 23 *S. marcescens* isolates of clinical and insect origin. A plus (+) indicates lysis & a minus (-) indicates no lysis. Spot assays performed in triplicate.

		Phage			
		Smarc1	**Smarc2**	**Smarc3**	**Smarc4**
Clinical *S. marcescens* Strains	**Y21 (Smarc1 Host)**	+	+	-	+
	Mutant 2	-	+	-	+
	Mutant 3	-	+	-	+
	Mutant 4	-	+	-	+
	G10 (Smarc2 Host)	+	+	-	+
	B61 (Smarc3 Host)	+	+	+	-
	Mutant 1	-	+	-	-
	Mutant 9	-	+	-	-
	Y37 (Smarc4 Host)	+	+	-	+
	Mutant 2.2	+	-	-	-
	Mutant 2.3	+	-	-	-
	Mutant 23	+	-	-	-
Insect *S. marcescens* Strains	**AM923**	-	+	-	+
	17-177	-	+	+	+
	16-1099	-	-	-	-
	17-042	-	+	-	+
	17-249	-	+	-	+
	18-593	+	+	-	-
	19-QA	-	-	-	-
	19-015	-	+	-	+
	LH33P3	-	-	-	+
	BI12	-	+	-	+
	CENT2.13	-	-	-	+

4.4 Development of liquid infection assay workflow

4.4.1 Phage specific primer design

Direct qPCR allows the detection of change in phage abundance over time to be quantified. It also allows quantification of phage in pure suspension to ensure consistent concentrations are applied across experiments. Therefore, primers were designed specific to each phage (Table 2.3) based on the previously sequenced genomes of Smarc1 – 4, to amplify fragments of ~250 bp in length. Using direct qPCR, each primer was subsequently cross checked against each phage to ensure specificity. Each suspension returned copy number (cp) per µl values between ~$3x10^5$ – $7x10^6$ (Table 4.3). Each primer pair displayed specific binding. Primer pair 4 displayed nominal cp/µl values for Smarc1-3, likely resulting from primer dimerization.

Table 4.3 Primer specificity cross check. Primer pairs were challenged against each individual phage suspension to assess specificity. Copy number (cp) values represent the number of phage particles per µl of suspension.

		Primer Pair			
		1	**2**	**3**	**4**
Smarc	**1**	6.1 $x10^5$ cp/µl	0 cp/µl	0 cp/µl	25 cp/µl
	2	0 cp/µl	2.9 $x10^5$ cp/µl	0 cp/µl	13 cp/µl
	3	0 cp/µl	0 cp/µl	6.9 $x10^6$ cp/µl	14 cp/µl
	4	0 cp/µl	0 cp/µl	0 cp/µl	1.2 $x10^6$ cp/µl

4.4.2 qPCR phage enumeration

The multiplicity of infection (MOI) applied in the liquid infection assays needed to be controlled, as different MOI's have different effects on host growth. As such, phage concentration in their pure suspensions needed to be quantified. The established method of determining the concentration of phage in suspension is counting the number of plaques produced when a known quantity of pure phage suspension is dropped onto a bacterial lawn of its host, resulting in a plaque forming unit (PFU) per ml value. The workflow described in this section necessitates that the phage concentration in suspension is known immediately prior to the commencement of each of the liquid infection assays, as phage numbers can deteriorate in suspension at different rates from each other. To expediate this time-consuming process, phage enumeration was performed as part of the workflow via direct qPCR, resulting in a cp/ml value. To ensure the reliability of this method, the PFU/ml and cp/ml values of the Smarc suspensions were compared (Table 4.4).

Table 4.4 Comparison of phage enumeration techniques. Enumeration of phage particles in pure suspension via direct qPCR (cp/ml values) and plaque count (PFU/ml values). The cp/PFU ratios indicate how comparable the two results are. Values represent the mean of *n*=3.

Smarc	cp/ml	PFU/ml	cp/PFU
1	2.5×10^8	2.1×10^9	0.1
2	5.4×10^9	3.0×10^9	1.8
3	1.1×10^{11}	1.0×10^{11}	1.1
4	2.5×10^8	2.0×10^8	1.3

4.4.2 Phage cocktail design and formulation

To demonstrate the efficacy of phage cocktails at suppressing WT and mutant *S. marcescens* growth, cocktails were designed based on the Smarc host range data (Table 4.1). Cocktails 1 & 2 (CTs 1 & 2) contained three phages while CT's 3 & 4 contained two. Smarc1 & 2 cause lysis of each WT strain and were included in CT's 1 & 2. As Smarc3 lyses strain B61 but not strain G10 it was included in CT1 (Smarc1, 2 & 3). Conversely, Smarc4 lyses G10 but not B61 and was included in CT2 (Smarc1, 2 & 4). This meant that for each liquid infection assay, only phages shown to cause lysis of the WT strain on solid media were included in the applied cocktails. The remaining cocktails, CT3 (Smarc1 & 2) and CT4 (Smarc2 & 4) were designed to exclude the phage responsible for the original mutant resistant phenotype.

To ensure that consistent amounts of phage were applied across experiments, pure phage suspensions were quantified via qPCR immediately before the liquid infection assays. Following quantification, the suspensions were diluted to ~$1x10^5$ cp/µl and mixed in equal quantities to a final volume of 6ml. After mixing, the cocktails constituents were quantified to ensure evenly balanced mixtures (Figure 4.1). The concentration of each phage within each cocktail was within one order of magnitude; the greatest difference being Smarc3 in CT1 being ~3.5x less than Smarc1 & 2. The overall concentrations across the cocktails were within one order of magnitude.

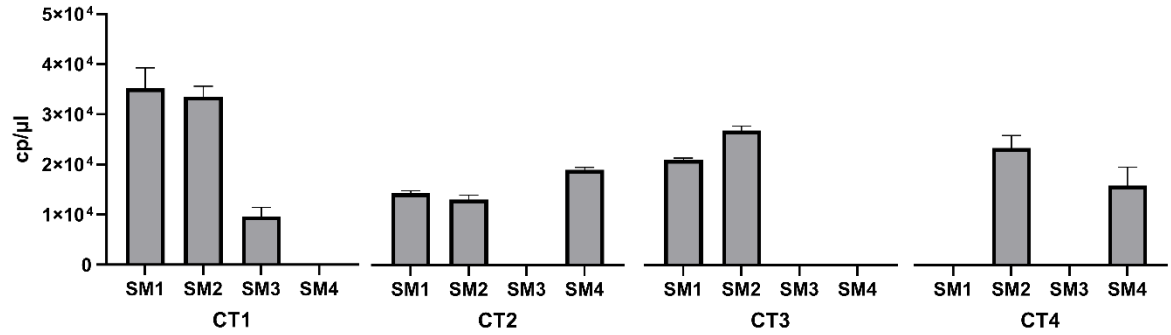

Figure 4.1 Concentration variability in cocktail constituents. Quantification of each phage within the four cocktails was performed to ensure even balances within the mixtures. Bars represent standard error of the mean (*n*=2).

4.4.3 Preliminary growth curve analysis and optimisation

To ensure comparability of results, preliminary growth curve analysis was undertaken of the *S. marcescens* strains Y21, G10, B61 & Y37 and their phage resistant mutants (Figure 4.2 a, b, c & d). OD_{600nm} readings were recorded using the CLARIOstar plate reader platform for 72 hours, allowing the growth of each mutant strain to be compared to their corresponding WT strain. WT and mutant strains experienced exponential growth for 8 – 10 hours, peaking at OD_{600nm} = ~1.7 – 1.9, before plateauing and entering stationary phase after ~18 hours. Subsequently, stable declines in OD_{600nm} readings were recorded for the remainder of the experiment. Differences between WT and mutant growth were minimal, with mutant growth reaching slightly lower OD_{600nm} readings, particularly in Y37 mutant 2.3.

To determine which MOI would be most effective at suppressing bacterial growth, CT1 was applied to Y21 mutant 2.1 at varying concentrations (Figure 4.2e). This example was chosen as it most clearly represented the effect increasing MOI had on host growth. At an MOI of 0.1, a slight reduction in OD_{600nm} was recorded throughout the experiment. An MOI of 1.0 had a more pronounced and sustained effect, whereas at an MOI of 10 host growth suffered the greatest initial suppression, before bacterial outgrowth was observed at ~26 hours. As an MOI of 10 had the most substantial initial killing effect, this concentration was set for all subsequent liquid infection assays.

Finally, to ensure the growth dynamics of the host/phage coculture were captured in their entirety, a preliminary liquid infection assay was performed for 88 hours (Figure 4.2f). Y21 was treated with three individual phages (each at MOI = 10) and two cocktails (each with a cumulative MOI of 10). For each treatment, an initial reduction in bacterial abundance was observed, which preceded bacterial outgrowth, replenishing bacterial abundance to nearly the same level as the Y21 monoculture. After ~36 hours, culture stabilisation was reached for all treatments and OD_{600nm} measurements steadily declined. As such, the incubation time for all subsequent liquid infection assays were set to 36 hours.

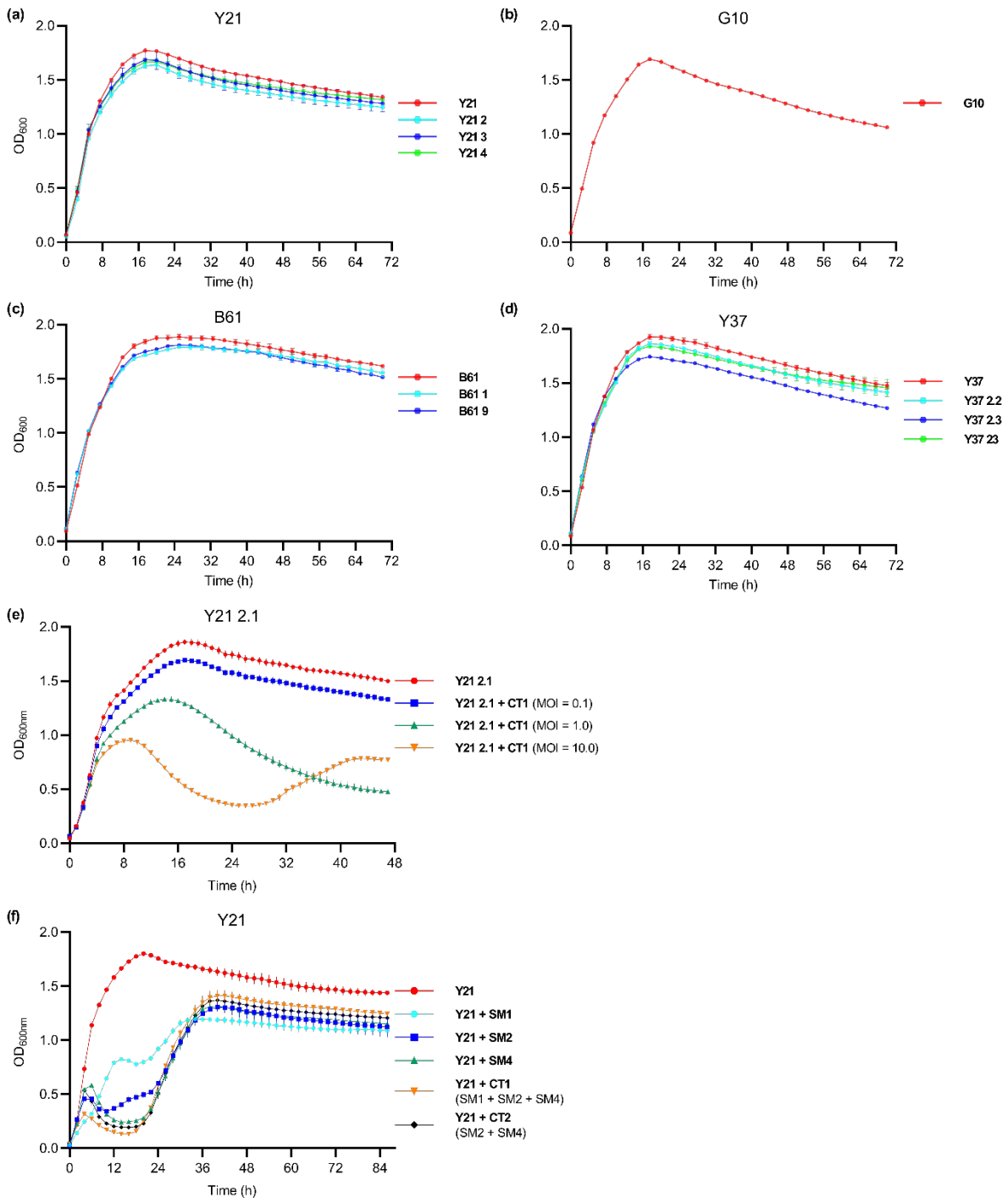

Figure 4.2 Preliminary growth curve analysis and assay optimisation. Growth curves of *S. marcescens* strains **(a)** Y21, **(b)** G10, **(c)** B61 & **(d)** Y37 and their corresponding phage resistant mutants incubated for 72 hours. **(e)** Sensitivity of *S. marcescens* strain Y21 2.1 to cocktail 1 (Smarc1, 2 & 4) at MOI's of 0.1, 1.0 & 10 measured over 48 hours. **(f)** Sensitivity of Y21 to three phages (Smarc1, 2 & 4) and cocktail 2 (Smarc1, 2 & 4) & cocktail 4 (Smarc2 & 4) was measured in a liquid infection assay for 88 hours to determine time required to reach culture stability. All incubations were set at 28° C. Bars represent standard error of the mean (*n*=3).

4.5 Liquid infection assays

Following the design and optimisation of the liquid infection assay workflow, the bacterial suppression capacity of the four Smarc phages and their cocktails was measured against the four *S. marcescens* host strains (Y21, G10, B61 & Y37) and their phage resistant mutants. For each WT host, their corresponding phage resistant mutants were assayed simultaneously and under identical conditions to ensure comparability of results. Planktonic monocultures of each host were spiked with three individual phages and two cocktails to an MOI of 10. The lytic ability of the phages was assessed as the reduction in OD_{600nm} readings over 36 hours as compared to the host monoculture (red circles, Figures 4.3 – 4.6). In all cases the individual phages caused suppression of their WT host, before bacterial re-growth occurred post 6 – 12 hours, replenishing bacterial abundance to varying degrees. As expected, the mutant strains persisted when treated with the phage responsible for their genesis, whilst being suppressed in some cases by the other individual phages. The phage resistant phenotype did not always confer complete resistance, with strains B61 mutants 1 & 9 (Figure 4.5 b, c) and Y37 mutants 2.3 & 23 (Figure 4.6 c, d) suffering minor sustained reductions in bacterial abundance whilst maintaining the characteristic growth curve of the untreated host.

The performance of the cocktails compared to their individual constituents can be categorised into three types; synergism (an additive killing effect is observed, where the cocktail is more effective than its constituents), neutrality (the cocktail performs as well as its most effective constituent) and antagonism (a reduction in phage lytic ability is observed when they are applied as a cocktail). Synergism, as defined for this study, might indicate that infection of the host by one phage in some way facilitates infection by another. This would enhance the overall lytic activity of the cocktail and elevate it above the sum of its parts. The term neutrality is used here to define a scenario where a cocktail of two or more phages displays no more lytic activity than its most effective constituent. This may arise when multiple phages are targeting the same receptor and one of them is simply outcompeting the others. The term antagonism, as it is used here, may suggest that the phages in the cocktail are somehow interfering with each other.

As an indication of which individual phages performed best in the cocktail applications, each cocktail & host coculture was manually sampled and subjected to direct qPCR at the beginning and end of each assay. This allowed the total change in abundance of each individual phage in the cocktail treated cultures to be measured. This data is represented as percentage values, listed

next to the cocktail constituents in the figure legends (Figure 4.3 – 4.6), and is summarised in table 4.5.

Table 4.5 Changes in phage abundance in liquid infection assay cocktail applications. Manual sampling at the beginning and end of each liquid infection assay, followed by direct qPCR, allowed quantification of the change in individual phage abundance when applied to each host as a cocktail. 100 µl of phage cocktail was added to each liquid culture at a concentration of ~$1x10^8$ cp/ml. Percentage values are shaded green (increase) or red (decrease) and serve as an indication of the performance of each phage in the cocktail. A dash (-) indicates no test.

		CT1			CT2			CT3		CT4	
		SM1	SM2	SM3	SM1	SM2	SM4	SM1	SM2	SM2	SM4
Y21	WT	-	-	-	-80%	-97%	14,477%	-	-	-98%	5,536%
	2.1	-	-	-	-73%	-55%	37,880%	-	-	-38%	37,660%
	3	-	-	-	-74%	-89%	65,641%	-	-	-88%	23,777%
	4	-	-	-	-69%	-90%	33,972%	-	-	-83%	3,926%
G10	WT	-	-	-	1,233%	-92%	-92%	740%	-92%	-	-
B61	WT	-47%	-98%	476%	-	-	-	-96%	-100%	-	-
	1	-85%	-98%	465%	-	-	-	-91%	-98%	-	-
	9	-90%	-98%	47%	-	-	-	-94%	-99%	-	-
Y37	WT	-	-	-	-33%	-99%	37,985%	-94%	-100%	-	-
	2.2	-	-	-	1,033%	-99%	1,442%	348%	-99%	-	-
	2.3	-	-	-	-95%	-100%	85%	-97%	-100%	-	-
	23	-	-	-	107%	-99%	20,830%	-62%	-100%	-	-

Strain Y21's (Smarc1 host) phage resistant mutants (strains 2.1, 3, 4) grew uninhibited under Smarc1 treatment. When treated with the other individual phages and phage cocktails they responded similarly to the WT strain (Figure 4.3). Cocktail treatment of Y21 and its mutants revealed no synergistic effects. CT2 experienced minor antagonism when challenged against Y21 WT and resistant strains 3 & 4 (Figure 4.3 a, c, d) with the individual Smarc4 treatment being more effective than the cocktail treatments. Despite causing lysis in the individual treatments, Smarc2 DNA copy numbers declined over the course of the experiment when applied as part of CT2 and CT4. Conversely, Smarc4 numbers increased for each of the cocktail treatments. This was true even for Y21, with Smarc1 numbers declining in the CT2 application despite Y21 being permissive for productive lysis in the individual treatment.

Strain G10 (Smarc2 host) produced no resistant mutants on solid media using the methods described in section 4.2. Despite this, bacterial outgrowth was observed within 2 hours of Smarc4 treatment, and after 6 hours of individual Smarc1 & 2 treatment (Figure 4.5). A synergistic effect was observed in the cocktail applications, with both CT2 and CT3 completely suppressing growth for the first 12 hours, before bacterial re-growth occurred. In both cocktail applications, an overall increase in abundance was recorded for Smarc1 only, while decreases were recorded for Smarc2 & 4. Re-growth for all treatments replenished bacterial abundance to ~62% of untreated levels.

CT1 had a slight synergistic effect against strain B61 (Smarc3 host), suppressing bacterial growth for 6 hours longer than CT2 and Smarc2 (Figure 4.5a). For the B61 mutants 1 & 9, CT1 and 3 exhibited the same effect as Smarc2 alone, indicating a lack of synergism against these strains (Figure 4.5 b, c). Smarc3 had little immediate effect against the resistant mutants, however by the end of the experiment it had reduced bacterial load by the same amount as the cocktails. As the cocktail treatment growth curves closely reflected the growth curves of the Smarc2 individual treatments, it could be expected that Smarc2 was responsible for the lytic effect observed in these treatments. However, only Smarc3 showed an increase in overall abundance in the CT1 application. Interestingly, in the CT3 application (containing only Smarc1 & 2), neither Smarc1 nor Smarc2 displayed an overall increase in DNA copy number.

Individual Smarc4 application against its host, strain Y37, resulted in significant and sustained suppression, whereas its phage resistant mutants remained largely unaffected (Figure 4.6). Smarc2 caused an initial suppression for all strains that preceded re-growth after ~6 hours.

Interestingly, Smarc1 application resulted in different growth curve characteristics for each of the mutants. CT2 displayed its most pronounced synergistic effects against the Y37 WT and resistant strains 2.2 & 23. Smarc4 replicated most against Y37 WT, according to the qPCR data, whereas Smarc1 appears to have contributed to the synergistic effect of CT2 against strains 2.3 & 23. Like the B61 CT3 application, Smarc1 & 2 showed an overall reduction in DNA copy number against the Y37 WT and resistant strains 2.3 & 23. However, a 348% increase was recorded in Smarc1 abundance in the CT3 application against mutant 2.2.

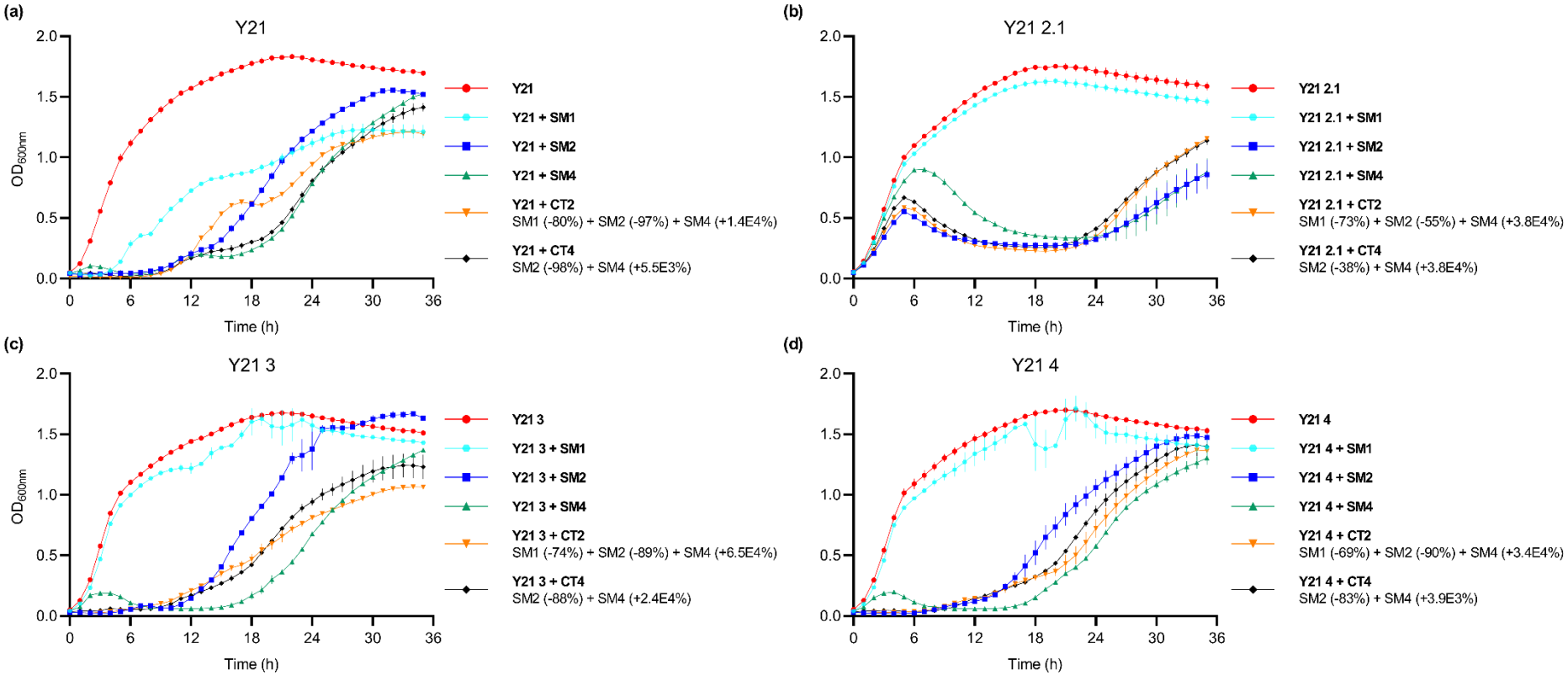

Figure 4.3 Sensitivity of Y21 WT (Smarc1 host) and selected Smarc1 resistant mutants to individual phages and phage cocktails. Sensitivity of **(a)** Y21 WT, **(b)** Y21 mutant 2.1, **(c)** Y21 mutant 3 & **(d)** Y21 mutant 4 to three phages (Smarc1, Smarc2 & Smarc4) and cocktail 2 (Smarc1 + Smarc2 + Smarc4) & cocktail 4 (Smarc2 + Smarc4) was measured in a liquid infection assay for 36 hours at 28° C. Individual phages were added to an MOI of 10 and cocktails were mixed in equal measure and added to a cumilative MOI of 10. Percentage values represent the percentage change in total abundance of each phage in the cocktail over the course of the experiment as measured by direct qPCR. Bars represent standard error of the mean (*n*=2).

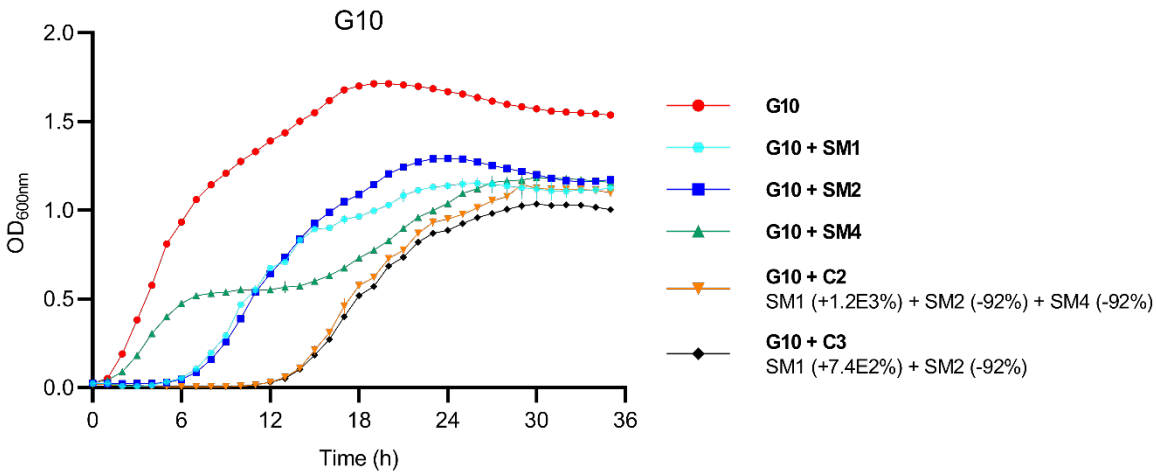

Figure 4.4 Sensitivity of G10 (Smarc2 host) to individual phages and phage cocktails. Sensitivity of G10 WT to three phages (Smarc1, Smarc2 & Smarc4) and cocktail 2 (Smarc1 + Smarc2 + Smarc4) & cocktail 3 (Smarc1 + Smarc2) was measured in a liquid infection assay for 36 hours at 28° C. Individual phages were added to an MOI of 10 and cocktails were mixed in equal measure and added to a cumilative MOI of 10. Percentage values represent the percentage change in total abundance of each phage in the cocktail over the course of the experiment as measured by direct qPCR. Bars represent standard error of the mean (*n*=2).

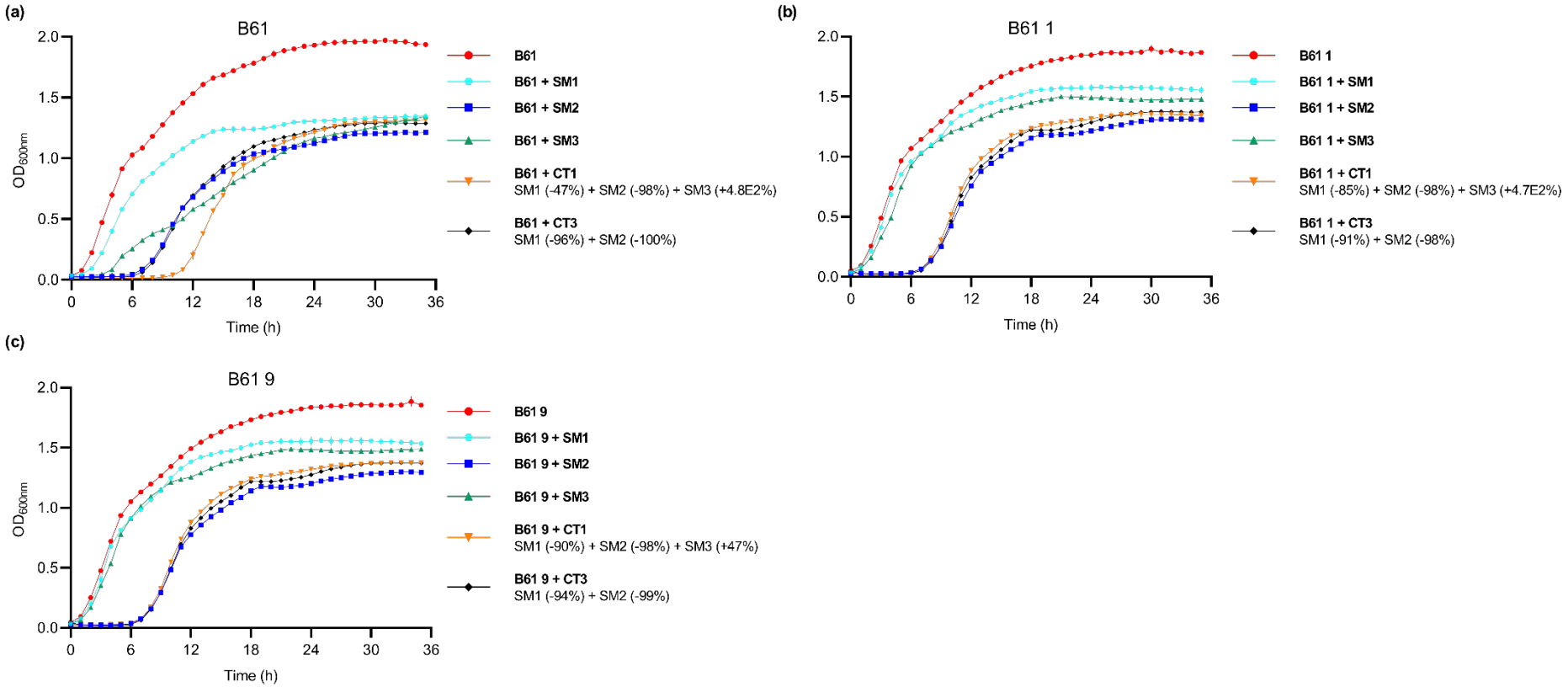

Figure 4.5 Sensitivity of B61 WT (Smarc3 host) and selected Smarc3 resistant mutants to individual phages and phage cocktails. Sensitivity of **(a)** B61 WT, **(b)** B61 mutant 1 & **(c)** B61 mutant 9 to three phages (Smarc1, Smarc2 & Smarc3) and cocktail 1 (Smarc1 + Smarc2 + Smarc3) & cocktail 3 (Smarc1 + Smarc2) was measured in a liquid infection assay for 36 hours at 28° C. Individual phages were added to an MOI of 10 and cocktails were mixed in equal measure and added to a cumilative MOI of 10. Percentage values represent the percentage change in total abundance of each phage in the cocktail over the course of the experiment as measured by direct qPCR. Bars represent standard error of the mean (n=2).

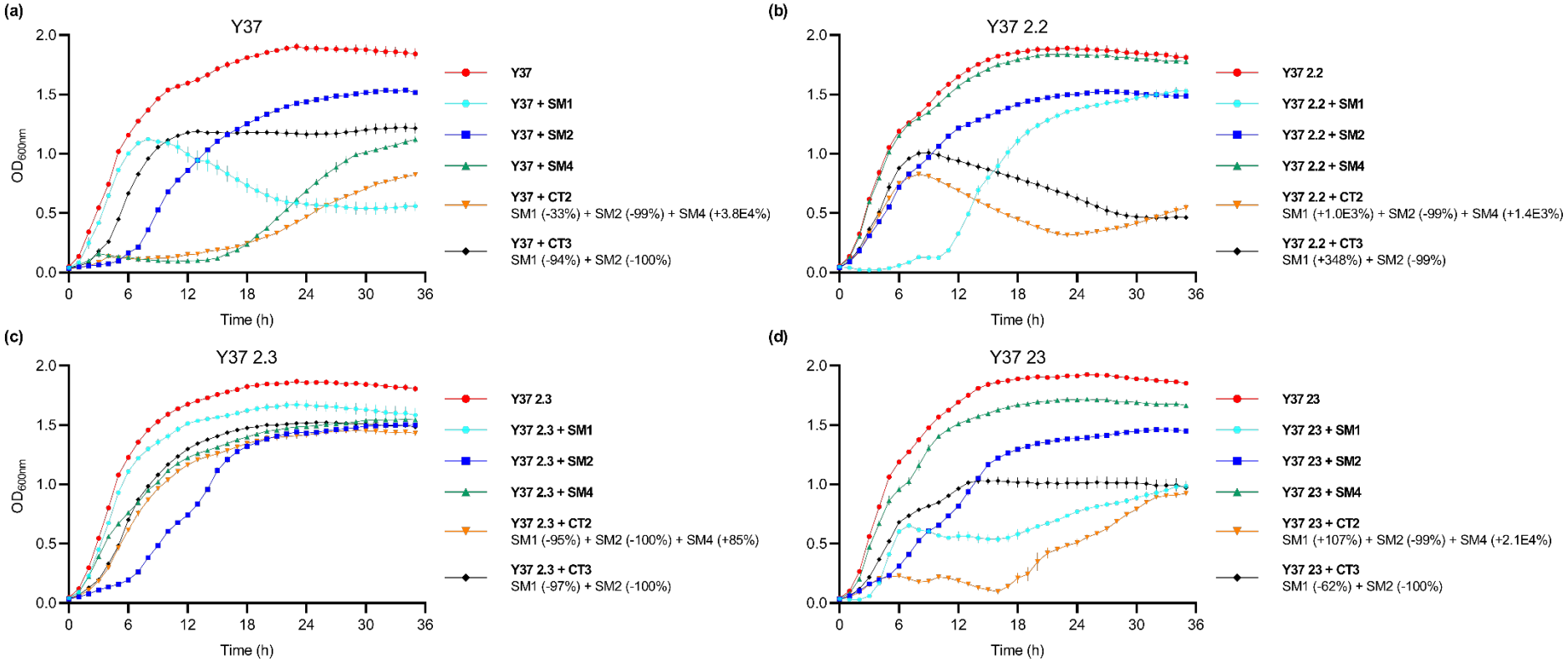

Figure 4.6 Sensitivity of Y37 WT (Smarc4 host) and selected Smarc4 resistant mutants to individual phages and phage cocktails. Sensitivity of **(a)** Y37 WT, **(b)** Y37 mutant 2.2, **(c)** Y37 mutant 2.3 & **(d)** Y37 mutant 23 to three phages (Smarc1, Smarc2 & Smarc4) and cocktail 2 (Smarc1 + Smarc2 + Smarc4) & cocktail 3 (Smarc1 + Smarc2) was measured in a liquid infection assay for 36 hours at 28° C. Individual phages were added to an MOI of 10 and cocktails were mixed in equal measure and added to a cumilative MOI of 10. Percentage values represent the percentage change in total abundance of each phage in the cocktail over the course of the experiment as measured by direct qPCR. Bars represent standard error of the mean (*n*=2).

Table 4.6 Summary of cocktail effects based on liquid infection assays. S = synergism (an additive killing effect is observed, where the cocktail is more effective than its constituents), N = neutral (the cocktail performs as well as its most effective constituent) & A = antagonistic (a reduction in phage lytic ability is observed when they are applied as a cocktail). A dash (-) indicates no test. The definitions of synergism, neutrality, and antagonism defined here pertain to this study only; their definitions are discussed on p. 54.

	Y21				G10	B61			Y37			
	WT	**2.1**	**3**	**4**	**WT**	**WT**	**1**	**9**	**WT**	**2.2**	**2.3**	**23**
CT1	-	-	-	-	-	S	N	N	-	-	-	-
CT2	A	N	A	A	S	-	-	-	S	S	A	S
CT3	-	-	-	-	S	N	N	N	A	S	N	N
CT4	N	N	N	N	-	-	-	-	-	-	-	-

CT1 was suitable for testing against the smallest number of strains, displaying synergism against the B61 WT and neutrality against its two mutants (Table 4.6). CT2 displayed predominantly synergistic effects against the strains tested; it performed synergistically against G10 and all Y37 strains except mutant 2.3, however it had an antagonistic effect against all Y21 strains except 2.1 for which it was neutral. CT3 was synergistic against G10 and Y37 mutant 2.2, neutral against all B61 strains and Y37 mutants 2.3 & 23, and antagonistic against the Y37 WT. CT4 was only tested against the Y21 strains and performed neutrally compared to the individual phages.

Section 5: Discussion

5.1 Introduction

There were two main objectives of this study. The first aim was to phylogenetically characterise 32 *Serratia* phage genomes, including the six novel *S. marcescens* phages isolated and sequenced by our lab. A contemporary framework was devised for clustering phage genomes based on similarity at the nucleotide and protein coding levels, and genome synteny. This analysis places the phages in context amongst the global phage sequence space and provides a reference point for characterisation of more novel *S. marcescens* phages as they are discovered.

The second aim was to begin the *in vitro* characterisation of the Smarcs capable of lysing our panel of clinical *S. marcescens* isolates in liquid culture. This involved assessing their ability to suppress the growth of WT and phage resistant mutant strains, either as individual phages or mixed into a cocktail. The emergence of synergistic and antagonistic effects within the cocktail applications was monitored, and the individual performance of the phages in these cocktail applications was quantified.

5.2 Clustering the *Serratia* phages

The function of genome clustering is to group related phages in a way that is biologically significant. This can facilitate investigation into their common traits and genetic diversity. Clustering is typically carried out by assessing phages overall nucleotide identity and gene content similarity. An important element of genome clustering is marrying established phylogenetic determination with contemporary approaches that can effectively group phages despite the complications introduced by HGT. Cluster, genus and species criteria continue to be refined as phage biology is better understood and technology affords more effective analysis (69). As such, the nature of genome clustering is not arbitrary, but it does vary among researchers. This reflects the difficulty of assigning quantifiers that meaningfully represent the biological relationships that exist within the phage virome.

5.2.1 Dot plot clusters

The criteria used in this work resulted in 20 phages being assigned to 8 clusters, with the remaining 12 not sharing significant nucleotide sequence similarity to any known phage (Table 3.3). Grouping into clusters using various agreeable methods lends surety to the biological significance of these results. In this case, the patterns of similarity observable through dot plot analysis are supported by genome alignment maps and intergenomic sequence similarity calculation.

The dot plot analysis performed here highlights how stretches of homology are interrupted by periods of dissimilarity, arising from the modular nature of the *Serratia* phage genomes (Figure 3.4). These modules are likely formed by the recombination of host genomic material into phage genomes during the replicative process (68). In the event of simultaneous infection, it is also possible for a phage to incorporate genomic material from other phages into its own genome (68). The short stretches of dissimilarity that flank homologous regions in Fig 3.8., may have resulted from the introduction of novel genes through HGT, or excision of unnecessary genes that would inhibit efficient genome packaging into the viral capsid (68).

Pairwise comparison of the cluster C phages (Smarc 5 & 6) show that regions of sequence homology are located at similar points along the genomes, indicating conservation of synteny (Figure 3.4a). This is especially visible in the genome map of the Cluster C phages in Fig. 3.8.

Dot plots reveal that this pattern of synteny is conserved across nearly all of the clusters (Figure 3.4b) supporting that these are genetically related lineages (110). Where gene order is disrupted, as in Cluster D, switching of large multi-gene modules is observed rather than random shuffling of the genome. As gene order is functional for a number of gene cassettes, it is sensical that genetic selection favours genomes that conserve the established gene order (110). Non-homologous recombination that disrupts gene order may result in inviable progeny that fail to propagate within the population (110).

The genetic distance between clustered phages, as determined by the Neighbour-Joining tree build method based on alignment of the *Serratia* phages LTS genes, is short relative to those between clusters (Figure 3.6). Whilst this approach is insensitive to the effect HGT has on phage genomes, it does provide a statistically sound inference of phylogenetic distance. The dot plot cluster designations agree with this phylogenetic assessment, falling within the clades of the tree. LTS sequences are present within 31 of the 32 *Serratia* phages, however they are not ubiquitous in phage sequence space, which limits their usefulness when comparing distantly related phages. Gene content network analysis can give context to a phages position in the global sequence space, where marker gene phylogeny cannot (68, 72).

5.2.2 Network phylogeny

VCs identified by vConTACT2 are based on gene content similarity between phages (85). As this type of phylogeny is not restricted by the necessity of devising individual lineages, it is capable of expressing a reticulated representation of phage relatedness. Gene sharing networks do not attempt to assign shared ancestors, and the distances between nodes cannot be directly correlated with genetic distance. Instead, they place phages in context among the global sequence space, illustrating their interrelatedness and tendency to exchange genetic material.

Fig 3.1 & 3.2 reveal that gene sharing is restricted by host preference, but not as much as it is by morphology. The rate that phages exchange genes within different host groups may be different, as illustrated by the varying levels of genetic isolation observable in the network. The *Serratia, Pseudomonas, Salmonella* and *Escherichia* phage VCs within the Proteobacteria supercluster appear more muddled and interconnected than the *Mycobacterium* and *Gordonia* phage VCs within their supercluster. This could indicate that phages infecting hosts of the Proteobacteria phylum are more promiscuous, and perhaps the genes required to infect these

hosts are more easily interchangeable. Whilst this conjecture needs to be validated experimentally, it represents the potential inferences that could be made using gene sharing networks like these. Such an investigation would be aided by the creation of a bipartite network that not only illustrates the relationship between phages, but also which genes are similar among them (105).

Finally, where the cluster designations are overlayed onto the gene sharing network (Figure 3.7), we can see there is good agreement with the VCs. This indicates that the vConTACT2 VCs are supported by whole genome nucleotide similarity assessment and marker gene phylogeny. Interestingly, some of the singletons inhabit VCs belonging to clustered phages. The fact that they were not clustered by nucleotide similarity assessment may be due to the high levels of sequence divergence characteristic of phages. It's possible that evolutionary pressure to maintain protein function facilitates the conservation of coding sequence similarity, while nucleotide similarity is lost (68).

5.2.3 Smarc taxonomy

The 6 Smarcs had been previously assigned to the family *Siphoviridae* based on morphological characterisation via transmission electron microscopy (Petrovski, S. pers. comm.). 46 Smarc related phages, identified by the gene sharing network (Figure 3.1), were assessed to determine if any of the Smarcs could be putatively assigned to a species. As >95% nucleotide sequence identity is required for strains to be considered the same species, ICTV's criteria is not met by any of the Smarcs (69). However, Smarc 5 & 6 both share >50% intergenomic sequence similarity to a group of *Salmonella* phages of the genus *Chivirus*. This speaks to the relative liberty with which phages infecting different Proteobacteria hosts appear to share genes.

5.3 *In vitro* characterisation of the Smarc phages

Following phylogenetic characterisation of Smarc1-6, the phages suitability for therapeutic use was tested *in vitro.* Previous studies have demonstrated the potential for *S. marcescens* phages to be used therapeutically, either by themselves or in combination with antibiotics (82). Despite phage cocktails showing potential to elevate the lytic activity of their constituents and discourage resistant bacterial outgrowth, no phage cocktail study has been published regarding *S. marcescens* phages. In this report, we measure the direct killing effect of the phages and their cocktails against WT and phage resistant *S. marcescens* strains.

Some phages capable of lysing their host on solid media are unable to productively lyse in planktonic culture (82, 111). Smarc5 & 6 did not cause a reduction in OD_{600nm} when applied to liquid cultures of their respective hosts. Future application of these phages may require lytic ability in liquid suspension; as such, only Smarc1-4 were selected for further testing against a panel of clinical *S. marcescens* isolates.

5.3.1 Smarc host ranges

Host range testing serves as an initial indication of therapeutic potential. Broad host range phages are desirable in cocktail applications as they broaden the lytic ability of the mixture in clinical settings where the susceptibility of an infective agent to individual phages is unknown (112). Smarc2 & 4 demonstrated the broadest lysing ability against the strains tested (Table 4.2), suggesting they might be the most useful in clinical and industrial settings. Smarc3 displayed the narrowest lytic range, however, it was the only phage to lyse clinical *S. marcescens* B61 on solid media. This suggests that a narrow host range phage such as Smarc3 could still be important in reaching clinical outcomes.

The 11 insect *S. marcescens* strains supplied to us for host range testing were isolated by the U-Vet Werribee Animal Hospital (University of Melbourne) from a captive collection of Lord Howe Island (LHI) Stick Insects. *S. marcescens* has been isolated from several mortalities within the population and is thought to have contributed to their fate (Marenda, M pers. comm.). The strains targeted in the host range analysis were sequenced and selected for genetic diversity (Marenda, M pers. comm.), however no comparative analyses of the strains were performed in this work. Annotation of the host genomes would allow identification of

prophages that may be inhibiting phage superinfection. While these strains were not included in the subsequent liquid infection assays, they increased the host range testing sample size and supported that Smarc2 & 4 have the broadest lytic activity.

The four clinical *S. marcescens* strains used in this study were isolated from patients at the Royal Melbourne Hospital. Following their isolation, phage resistant mutants were generated for strains Y21, B61 & Y37 *in vitro.* These mutants underwent WGS to identify the type and location of the mutation responsible for their phage resistant phenotype (Table 4.1). Many resistant mutants were generated; however, strains were selected so that each mutant represented a different mutation location, i.e., no two strains in this study contained mutations in the same gene. Selecting for a diverse range of mutations in this way provides a panel of resistant phenotypes that may be more representative of *in vivo* scenarios. Host range analysis of the mutant strains was performed to monitor the development of cross-resistance to the phages; however, none was detected. This host range data served as the basis for the liquid infection assay experimental design.

5.3.2 Liquid infection assays

In vitro characterisation of the Smarc phages via liquid infection assays necessitated a few steps; the first being molecular enumeration of phage copy number in suspension. The plaque count method is a time-consuming procedure as it requires a bacterial lawn of the host to be incubated overnight. Additionally, environmental influences such as temperature can alter gene expression in the host and create inconsistent results. Described previously, molecular quantification of phage genomes, via direct qPCR, provides a rapid alternative to the plaque count method (113, 114). Duyvejonck *et al.,* note that the difference in yield between the two methods was phage dependant; as such, a correction factor could be established for each phage that accounted for this difference (113). A comparison of the two methods was performed for the Smarc phages (Table 4.4); Smarc2, 3 & 4 produced almost identical values. Interestingly, the number of plaques produced by Smarc1 exceeded the number of genome copies measured via qPCR. This may have resulted from imperfect qPCR parameters including annealing temperature, $MgCl_2$ concentration and primer specificity. Consideration of Smarc1's correction factor was carried forth to account for this. Quantification of the Smarc suspensions allowed consistent amounts of phage to be applied across the liquid infection assays. This was crucial as the amount of phage applied drastically affected the growth curve characteristics of the host

(Figure 4.2e). Applying inconsistent amounts would mean that no comparisons could be drawn between the lytic activity of the individual phages.

Regardless of the initial killing effect displayed by any of the phages, bacterial outgrowth was observed for each of the clinical *S. marcescens* strains. This was true for the cocktail applications too; however, in some cases the two and three phage cocktails produced a more prolonged suppression of bacterial growth (Figure 4.4 and Figure 4.6d respectively). One explanation for this synergistic effect is described by Schmerer *et al.,* who demonstrated that within a multi-phage system, one phage may augment the growth properties of another (115). A possible biological mechanism for this phenomenon is that excess phage enzymes, left over from progeny phage assembly, facilitate the degradation of bacterial extracellular structural components and improve accessibility for the other phage (115). Cocktail synergism resulting in prolonged bacterial suppression could have clinical benefits; one example is the enhanced window of action for augmenting therapies or host immunity to act against the target bacteria. This could result in complete eradication of a bacterial pathogen from a host system, protecting the host from the bacterial re-growth that is observed here.

The most interesting cocktail effect resulted from CT2 (Smarc1, 2 & 4) being applied to Y37 mutant 23 (Figure 4.6d). Despite being resistant to Smarc4, growth of Y37 mutant 23 was significantly more suppressed by the three-phage cocktail compared to CT3, which only included Smarc1 & 2. This suggests that the evolutionary pressure imposed by Smarc1 & 2 caused re-sensitisation of the host to Smarc4. This would be a desirable effect in clinical applications; however, it remains isolated to this one case. Further investigation to uncover the mechanism responsible for re-sensitisation would support our assessment of this observation. For example, adding additional qPCR sampling time points throughout the course of the experiment would allow identification of the re-sensitisation event, and molecular characterisation of the strain at this point could help determine if there was a genetic basis for re-sensitisation.

Quantification of the change in abundance of the individual phages during the cocktail applications identified conflicting results. Where synergism occurs, it may be expected that an increase in phage abundance is observed for two or more phages rather than one (115). This was true for Y37 mutants 2.2 and 23, with Smarc1 & 4 both having increased in number after 36 hours. On the other hand, CT2 & 3 displayed synergism against G10 but only Smarc1

registered an increase. Even more concerning was the complete lack of increase in phage numbers when CT3 (Smarc1 & 2) was applied against the B61 strains and Y37 WT, mutant 2.3 & 23 (Figure 4.5 & Figure 4.6a, c & d respectively). A possible explanation for this is contamination of CT3 with either Smarc3 (when testing the B61 strains) or Smarc4 (when testing the Y37 strains) at the time of application. Alternatively, phage numbers may have degraded in suspension by the time sampling occurred; the re-growth of bacteria suggests a lack of phage replication after 6 hours of incubation, providing time for phage DNA copy numbers to decline.

We have provided a broad and preliminary perspective of the potential effect the Smarcs will have against clinical *S. marcescens* isolates. By testing against a panel of clinical *S. marcescens* strains we have demonstrated that while some synergistic effects may be observed, phages mixed into a cocktail are not always more effective than their constituents. The effect of cocktail antagonism, as defined for this study, was nominal and unworthy of significant concern. This analysis provides a basis for further ecological examination that could help to elucidate the underlying mechanisms by which synergism occurs and phage resistance evolves.

5.4 Conclusion and future directions

Comparative genomic analysis revealed that 20 out of the 32 *Serratia* phage sequences could be clustered together with at least one other *Serratia* phage. The clustering criteria used in this study followed a framework previously devised by researchers in the field (116); similar studies have followed this trend (95, 96, 98). Incorporation of contemporary gene sharing networks into this framework illustrate the direction phage classification is heading in the face of an influx of phage genomic data (99).

Our approaches when phylogenetically characterising phages must evolve with our understanding of the mechanisms that shape their genomes. Nucleotide sequence similarity, genome synteny analysis and gene content clustering all support these as being genetically related lineages. This is, however, a narrow perspective of their traits and diversity. Phages demonstrate unique and varied morphologies, nucleic acid types, genome structures, lifestyles, and evolutionary histories. As such, the classifications used to describe phages need to be equally as sophisticated in order to capture their distinctive attributes. Phage clustering efforts like this have preceded investigation into the genetic diversity within the clusters and the core

genes that they share (116). Recent studies have investigated the determinants of horizontal gene transfer rates and categorized phages into low and high gene flux modes (66). As horizontal gene transfer rates within clusters of phages have been shown to be modulated by phage lifestyle (66), we can imagine that behavioural predictions could potentially be made from comparative analyses such as these. Ultimately, through continued sampling of the phage sequence space and further efforts to rationalise the phenomena observed within clusters, I imagine that inference about a phage's suitability for therapy will be able to be made purely from genomic data. One example might be selecting for broader host range phages based on their apparent tendency to share genes with varying host groups. This type of insight would be indispensable when trawling the environment for phages of clinical and industrial potential.

The therapeutic potential of four novel *S. marcescens* bacteriophages has been assessed with the aid of an optimised liquid infection assay workflow. Smarcs1 – 4 show promise as therapeutic agents, being free from undesirable genes and capable of causing sustained lysis of their hosts. Phage resistance presents a roadblock in the application of phages and is not circumvented here. Employing the Step-by-Step method may have resulted in more robust cocktails (78), where phages infecting the resistant mutant were isolated after the fact and included in the original cocktail, however time restrictions precluded this approach. Growing our armament of Smarcs using this targeted approach and adding additional characterisation methods, including assessment of synergism with antibiotics and host immunity (82, 117), will progress their journey towards clinical and industrial use. Additionally, the importance of phage receptor identification has become increasingly clear (74). Competition for shared phage receptors by multiple phages can thwart cocktail effectiveness, and may account for the synergistic or antagonistic effects observed here. Future efforts to elucidate the Smarcs receptors would aid cocktail design and may drastically enhance therapeutic prospects (74).

Section 6: References

1. **Falagas ME**, **Bliziotis IA**. Pandrug-resistant Gram-negative bacteria: the dawn of the post-antibiotic era? *International Journal of Antimicrobial Agents.* 2007;29(6):630-6.

2. **Arias CA, Murray BE.** Antibiotic-Resistant Bugs in the 21st Century — A Clinical Super-Challenge. *New England Journal of Medicine.* 2009;360(5):439-43.

3. **Wittebole X, De Roock S, Opal SM.** A historical overview of bacteriophage therapy as an alternative to antibiotics for the treatment of bacterial pathogens. *Virulence.* 2014;5(1):226-35.

4. **Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, et al.** Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *The Lancet Infectious Diseases.* 2018;18(3):318-27.

5. **Hejazi A, Falkiner FR.** *Serratia marcescens*. *Journal of Medical Microbiology.* 1997;46(11):903-12.

6. **Bulla Jr LA, Rhodes RA, St.** Julian G. Bacteria as insect pathogens. *Annual review of microbiology.* 1975;29(1):163-90.

7. **Steinhaus EA.** *Principles of Insect Pathology: Hafner Publishing Company; 1967.*

8. **vanEngelsdorp D, Hayes J, Jr., Underwood RM, Pettis J. A** Survey of Honey Bee Colony Losses in the U.S., Fall 2007 to Spring 2008. *PLOS ONE.* 2009;3(12):e4071.

9. **Smith KM, Loh EH, Rostal MK, Zambrana-Torrelio CM, Mendiola L, Daszak P.** Pathogens, Pests, and Economics: Drivers of Honey Bee Colony Declines and Losses. *EcoHealth.* 2013;10(4):434-45.

10. **Raymann K, Coon K, Shaffer Z, Salisbury S, Moran N.** Pathogenicity of *Serratia marcescens* Strains in Honey Bees. *mBio*. 2018;9.

11. **Burritt NL, Foss NJ, Neeno-Eckwall EC, Church JO, Hilger AM, Hildebrand JA, et al.** Sepsis and Hemocyte Loss in Honey Bees (*Apis mellifera*) Infected with *Serratia marcescens* Strain Sicaria. *PLOS ONE.* 2016;11(12):e0167752.

12. **El Sanousi SM, El Sarag MSA, Mohamed SE.** Properties of *Serratia marcescens* Isolated from Diseased Honeybee (*Apis mellifera*) Larvae. *Microbiology*. 1987;133(1):215-9.

13. **Doss J, Culbertson K, Hahn D, Camacho J, Barekzi N.** A Review of Phage Therapy against Bacterial Pathogens of Aquatic and Terrestrial Organisms. *Viruses*. 2017;9(3):50.

14. **Liu M, Gill J, Young R, Summer E.** Bacteriophages of wastewater foaming-associated filamentous *Gordonia* reduce host levels in raw activated sludge. *Scientific reports.* 2015;5:13754.

15. **Eaton MD, Bayne-Jones S.** Bacteriophage Therapy: Review Of The Principles And Results Of The Use Of Bacteriophage In The Treatment Of Infections. *Journal of the American Medical Association.* 1934;103(23):1769-76.

16. **Adeolu M, Alnajar S, Naushad S, S. Gupta R.** Genome-based phylogeny and taxonomy of the '*Enterobacteriales'*: proposal for *Enterobacterales ord. nov.* divided into the families *Enterobacteriaceae*, *Erwiniaceae fam. nov.*, *Pectobacteriaceae fam. nov.*, Yersiniac*eae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. International Journal of Systematic and Evolutionary Microbiology.* 2016;66(12):5575-99.

17. **Grimont F, Grimont P.** The Genus *Serratia*. 322006. p. 219-44.

18. **Marrie TJ, Costerton JW.** Prolonged survival of *Serratia marcescens* in chlorhexidine. *Applied and Environmental Microbiology.* 1981;42(6):1093.

19. **Nakashima AK, Highsmith AK, Martone WJ.** Survival of *Serratia marcescens* in benzalkonium chloride and in multiple-dose medication vials: relationship to epidemic septic arthritis. *J Clin Microbiol.* 1987;25(6):1019-21.

20. **Szewzyk U, Szewzyk R, Stenström TA.** Growth and survival of *Serratia marcescens* under aerobic and anaerobic conditions in the presence of materials from blood bags. *J Clin Microbiol.* 1993;31(7):1826.

21. **Bartolomeo B.** Lettera di Bartolomeo Bizio al chiarissimo canonico Angelo Bellani sopra il fenomeno della polenta porporina. *Biblioteca Italiana o sia Giornale di letteratura, scienze e arti.* 1823;30:275-95.

22. **Yu VL.** *Serratia marcescens. New England Journal of Medicine.* 1979;300(16):887-93.

23. **Wheat RP, Zuckerman A, Rantz LA.** Infection due to Chromobacteria: report of eleven cases. *AMA archives of internal medicine.* 1951;88(4):461-6.

24. **Jones RN.** Microbial Etiologies of Hospital-Acquired Bacterial Pneumonia and Ventilator-Associated Bacterial Pneumonia. *Clinical Infectious Diseases.* 2010;51(Supplement_1):S81-S7.

25. **van der Vorm ER.** Source, carriers, and management of a *Serratia marcescens* outbreak on a pulmonary unit. *Journal of Hospital Infection.* 2002;52(4):263-7.

26. **Kawecki D, Kwiatkowski A, Sawicka-Grzelak A, Durlik M, Paczek L, Chmura A, et al.** Urinary Tract Infections in the Early Posttransplant Period After Kidney Transplantation: Etiologic Agents and Their Susceptibility. *Transplantation Proceedings.* 2011;43(8):2991-3.

27. **Schaberg DR, Highsmith AK, Wachsmuth IK.** Resistance Plasmid Transfer by *Serratia marcescens* in Urine. *Antimicrobial Agents and Chemotherapy.* 1977;11(3):449.

28. **Enciso-Moreno JA, Pernas-Buitrón N, Ortiz-Herrera M, Coria-Jiménez R.** Identification of *Serratia marcescens* populations of nosocomial origin by RAPD-PCR. *Archives of Medical Research.* 2004;35(1):12-7.

29. **Pitt TL.** State of the art: typing of *Serratia marcescens*. *Journal of Hospital Infection.* 1982;3(1):9-14.

30. **Funke G, Monnet D, deBernardis C, von Graevenitz A, Freney J.** Evaluation of the VITEK 2 system for rapid identification of medically relevant gram-negative rods. *J Clin Microbiol.* 1998;36(7):1948-52.

31. **Rödel J, Mellmann A, Stein C, Alexi M, Kipp F, Edel B, et al.** Use of MALDI-TOF mass spectrometry to detect nosocomial outbreaks of *Serratia marcescens* and Citrobacter freundii. *Eur J Clin Microbiol Infect Dis.* 2019;38(3):581-91.

32. **Larose P, Picard B, Thibault M, Grimont F, Goullet P.** Nosocomial *Serratia marcescens* individualized by five typing methods in a regional hospital. *Journal of Hospital Infection.* 1990;15(2):167-72.

33. **Allen EG.** Conditions of the Colour Change of Prodigiosin. *Nature*. 1967;216(5118):929-31.

34. **Hines DA, Saurugger PN, Ihler GM, Benedik MJ.** Genetic analysis of extracellular proteins of *Serratia marcescens*. *Journal of Bacteriology.* 1988;170(9):4141.

35. **Alberti L, Harshey RM.** Differentiation of *Serratia marcescens* 274 into swimmer and swarmer cells. *Journal of Bacteriology.* 1990;172(8):4322.

36. **Matsuyama T, Fujita M, Yano I.** Wetting agent produced by *Serratia marcescens*. *FEMS Microbiology Letters.* 1985;28(1):125-9.

37. **Mahlen SD.** *Serratia* Infections: from Military Experiments to Current Practice. *Clinical Microbiology Reviews.* 2011;24(4):755.

38. **Wilson JW, Schurr MJ, LeBlanc CL, Ramamurthy R, Buchanan KL, Nickerson CA.** Mechanisms of bacterial pathogenicity. *Postgraduate Medical Journal.* 2002;78(918):216.

39. **Yamamoto T, Ariyoshi A, Amako K.** Fimbria-Mediated Adherence of *Serratia marcescens* Strain US5 to Human Urinary Bladder Surface. *Microbiology and Immunology.* 1985;29(7):677-81.

40. **Van Houdt R, Givskov M, Michiels CW.** Quorum sensing in *Serratia*. *FEMS Microbiology Reviews.* 2007;31(4):407-24.

41. **Abisado RG, Benomar S, Klaus JR, Dandekar AA, Chandler JR.** Bacterial Quorum Sensing and Microbial Community Interactions. *mBio*. 2018;9(3):e02331-17.

42. **Pinna A, Usai D, Sechi LA, Carta A, Zanetti S.** Detection of virulence factors in *Serratia* strains isolated from contact lens-associated corneal ulcers. *Acta Ophthalmologica.* 2011;89(4):382-7.

43. **Dart JKG, Stapleton F, Minassian D, Dart JKG.** Contact lenses and other risk factors in microbial keratitis. *The Lancet.* 1991;338(8768):650-3.

44. **Parment PA.** The role of *Serratia marcescens* in soft contact lens associated ocular infections. *Acta Ophthalmologica Scandinavica.* 1997;75(1):67-71.

45. **Crosa JH, Mey AR, Payne SM.** Iron transport in bacteria: ASM press Washington, DC; 2004.

46. **Motta EVS, Raymann K, Moran NA.** Glyphosate perturbs the gut microbiota of honey bees. *Proceedings of the National Academy of Sciences.* 2018;115(41):10305.

47. **Raymann K, Shaffer Z, Moran NA.** Antibiotic exposure perturbs the gut microbiota and elevates mortality in honeybees. *PLoS Biol.* 2017;15(3):e2001861.

48. **Sfakinos J.** Detecting the dual presence of AmpC and ESBL enzymes. *Microbiology Australia.* 2009;30(5):208-9.

49. **Jacoby GA.** AmpC beta-lactamases. *Clinical microbiology reviews.* 2009;22(1):161-82.

50. **Crivaro V, Bagattini M, Salza MF, Raimondi F, Rossano F, Triassi M, et al.** Risk factors for extended-spectrum β-lactamase-producing *Serratia marcescens* and *Klebsiella pneumoniae* acquisition in a neonatal intensive care unit. *Journal of Hospital Infection.* 2007;67(2):135-41.

51. **Codjoe FS, Donkor ES.** Carbapenem Resistance: A Review. *Med Sci (Basel).* 2017;6(1):1.

52. **Deshpande LM, Rhomberg PR, Sader HS, Jones RN.** Emergence of serine carbapenemases (KPC and SME) among clinical strains of *Enterobacteriaceae* isolated in the United States Medical Centers: Report from the MYSTIC Program (1999–2005). *Diagnostic Microbiology and Infectious Disease.* 2006;56(4):367-72.

53. **Cai JC, Zhou HW, Zhang R, Chen G-X.** Emergence of *Serratia marcescens*, Klebsiella *pneumoniae*, and *Escherichia coli* Isolates possessing the plasmid-mediated carbapenem-hydrolyzing beta-lactamase KPC-2 in intensive care units of a Chinese hospital. *Antimicrobial agents and chemotherapy.* 2008;52(6):2014-8.

54. **Shaw KJ, Rather PN, Sabatelli FJ, Mann P, Munayyer H, Mierzwa R, et al.** Characterization of the chromosomal aac(6')-Ic gene from *Serratia* marcescens. *Antimicrobial Agents and Chemotherapy.* 1992;36(7):1447.

55. **Hanlon GW.** Bacteriophages: an appraisal of their role in the treatment of bacterial infections. *International Journal of Antimicrobial Agents.* 2007;30(2):118-28.

56. **Weinbauer MG.** Ecology of prokaryotic viruses. *FEMS Microbiology Reviews.* 2004;28(2):127-81.

57. **d'Herelle F.** Bacteriophage as a Treatment in Acute Medical and Surgical Infections. *Bull N Y Acad Med.* 1931;7(5):329-48.

58. **Ackermann HW.** Bacteriophage taxonomy. *Microbiol Aust.* 2011;32(2):90-4.

59. **Ackermann HW.** Bacteriophage observations and evolution. *Research in Microbiology.* 2003;154(4):245-51.

60. **Barylski J, Kropinski AM, Alikhan N-F, Adriaenssens EM.** ICTV Virus Taxonomy Profile: *Herelleviridae*. *Journal of General Virology.* 2020.

61. **Adriaenssens EM, Wittmann J, Kuhn JH, Turner D, Sullivan MB, Dutilh BE, et al.** Taxonomy of prokaryotic viruses: 2017 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Archives of Virology.* 2018;163(4):1125-9.

62. **Kutter E, Sulakvelidze A.** Bacteriophages: biology and applications: Crc press; 2004.

63. **Sharma S, Chatterjee S, Datta S, Prasad R, Dubey D, Prasad RK, et al.** Bacteriophages and its applications: an overview. *Folia Microbiologica.* 2017;62(1):17-55.

64. **Seed KD.** Battling Phages: How Bacteria Defend against Viral Attack. *PLOS Pathogens.* 2015;11(6):e1004847.

65. **Chan BK, Abedon ST.** Chapter 1 - Phage Therapy Pharmacology: Phage Cocktails. In: Laskin AI, Sariaslani S, Gadd GM, editors. Advances in Applied Microbiology. *78: Academic Press; 2012.* p. 1-23.

66. **Mavrich TN, Hatfull GF.** Bacteriophage evolution differs by host, lifestyle and genome. *Nature Microbiology.* 2017;2(9):17112.

67. **Soucy SM, Huang J, Gogarten JP.** Horizontal gene transfer: building the web of life. *Nature Reviews Genetics.* 2015;16(8):472-82.

68. **Lawrence JG, Hatfull GF, Hendrix RW.** Imbroglios of Viral Taxonomy: Genetic Exchange and Failings of Phenetic Approaches. *Journal of Bacteriology.* 2002;184(17):4891-905.

69. **Barylski J, Enault F, Dutilh BE, Schuller MB, Edwards RA, Gillis A, et al.** Analysis of Spounaviruses as a Case Study for the Overdue Reclassification of Tailed Phages. *Systematic Biology.* 2019;69(1):110-23.

70. **Simmonds P, Adams MJ, Benkő M, Breitbart M, Brister JR, Carstens EB, et al.** Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology.* 2017;15(3):161-8.

71. **Ackermann H-W.** Tailed Bacteriophages: The Order Caudovirales. In: Maramorosch K, Murphy FA, Shatkin AJ, editors. Advances in Virus Research. 51: Academic Press; 1998. p. 135-201.

72. **Lima-Mendez G, Van Helden J, Toussaint A, Leplae R.** Reticulate Representation of Evolutionary and Functional Relationships between Phage Genomes. *Molecular Biology and Evolution.* 2008;25(4):762-77.

73. **Romero-Calle D, Guimarães Benevides R, Góes-Neto A, Billington C.** Bacteriophages as Alternatives to Antibiotics in Clinical Care. *Antibiotics (Basel).* 2019;8(3).

74. **Gordillo Altamirano FL, Barr JJ.** Unlocking the next generation of phage therapy: the key is in the receptors. *Current Opinion in Biotechnology.* 2021;68:115-23.

75. **Yen M, Cairns LS, Camilli A.** A cocktail of three virulent bacteriophages prevents Vibrio cholerae infection in animal models. *Nature Communications.* 2017;8(1):14187.

76. **Forti F, Roach DR, Cafora M, Pasini ME, Horner DS, Fiscarelli EV, et al.** Design of a Broad-Range Bacteriophage Cocktail That Reduces *Pseudomonas aeruginosa* Biofilms and Treats Acute Infections in Two Animal Models. *Antimicrobial Agents and Chemotherapy.* 2018;62(6):e02573-17.

77. **Bai J, Jeon B, Ryu S.** Effective inhibition of *Salmonella Typhimurium* in fresh produce by a phage cocktail targeting multiple host receptors. Food *Microbiology*. 2019;77:52-60.

78. **Gu J, Liu X, Li Y, Han W, Lei L, Yang Y, et al.** A Method for Generation Phage Cocktail with Great Therapeutic Potential. *PLOS ONE.* 2012;7(3):e31698.

79. **Flyg C, Xanthopoulos KG.** Insect Pathogenic Properties Of *Serratia Marcescens*. Passive And Active Resistance To Insect Immunity Studied With Protease-Deficient And Phage-Resistant Mutants. *Microbiology.* 1983;129(2):453-64.

80. **Yang Y, Shen W, Zhong Q, Chen Q, He X, Baker JL, et al.** Development of a Bacteriophage Cocktail to Constrain the Emergence of Phage-Resistant *Pseudomonas* aeruginosa. *Frontiers in Microbiology.* 2020;11(327).

81. **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*. *Nature Medicine.* 2019;25(5):730-3.

82. **Weber L, Jansen M, Krüttgen A, Buhl EM, Horz H-P.** Tackling Intrinsic Antibiotic Resistance in *Serratia Marcescens* with A Combination of Ampicillin/Sulbactam and Phage SALSA. *Antibiotics* (Basel). 2020;9(7):371.

83. **Matsushita K, Uchiyama J, Kato S-i, Ujihara T, Hoshiba H, Sugihara S, et al.** Morphological and genetic analysis of three bacteriophages of *Serratia marcescens* isolated from environmental water. *FEMS Microbiology Letters.* 2009;291(2):201-8.

84. **Savalia D, Westblade LF, Goel M, Florens L, Kemp P, Akulenko N, et al.** Genomic and proteomic analysis of phiEco32, a novel *Escherichia coli* bacteriophage. *J Mol Biol.* 2008;377(3):774-89.

85. **Dethlefsen L, Huse S, Sogin ML, Relman DA.** The Pervasive Effects of an Antibiotic on the Human Gut Microbiota, as Revealed by Deep 16S rRNA Sequencing. *PLoS Biol.* 2008;6(11):e280.

86. **Friman V-P, Buckling A.** Effects of predation on real-time host–parasite coevolutionary dynamics. *Ecology Letters.* 2013;16(1):39-46.

87. **Zhang J, Ketola T, Örmälä-Odegrip A-M, Mappes J, Laakso J.** Coincidental Loss of Bacterial Virulence in Multi-Enemy Microbial Communities. *PLOS ONE.* 2014;9(11):e111871.

88. **Voelker R.** FDA Approves Bacteriophage Trial. *JAMA*. 2019;321(7):638-.

89. **Chen L, Yuan S, Liu Q, Mai G, Yang J, Deng D, et al.** In Vitro Design and Evaluation of Phage Cocktails Against *Aeromonas salmonicida*. *Frontiers in Microbiology.* 2018;9(1476).

90. **Hesse S, Rajaure M, Wall E, Johnson J, Bliskovsky V, Gottesman S, et al.** Phage Resistance in Multidrug-Resistant *Klebsiella pneumoniae* ST258 Evolves via Diverse Mutations That Culminate in Impaired Adsorption. *mBio*. 2020;11(1):e02530-19.

91. **Nale JY, Vinner GK, Lopez VC, Thanki AM, Phothaworn P, Thiennimitr P, et al.** An Optimized Bacteriophage Cocktail Can Effectively Control *Salmonella* in vitro and in *Galleria mellonella*. *Frontiers in Microbiology.* 2021;11(3370).

92. **Korf IHE, Kittler S, Bierbrodt A, Mengden R, Rohde C, Rohde M, et al.** In Vitro Evaluation of a Phage Cocktail Controlling Infections with *Escherichia coli*. *Viruses*. 2020;12(12).

93. **Gatea Kaabi SA, Musafer HK.** New Phage cocktail against infantile Sepsis bacteria. *Microbial Pathogenesis.* 2020;148:104447.

94. **Kornienko M, Kuptsov N, Gorodnichev R, Bespiatykh D, Guliaev A, Letarova M, et al.** Contribution of *Podoviridae* and *Myoviridae* bacteriophages to the effectiveness of anti-staphylococcal therapeutic cocktails. *Scientific Reports.* 2020;10(1):18612.

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

96. **Gao R, Naushad S, Moineau S, Levesque R, Goodridge L, Ogunremi D.** Comparative genomic analysis of 142 bacteriophages infecting Salmonella enterica subsp. enterica. *BMC Genomics.* 2020;21(1):374.

97. **Szymczak P, Rau MH, Monteiro JM, Pinho MG, Filipe SR, Vogensen FK, et al.** A comparative genomics approach for identifying host-range determinants in *Streptococcus* thermophilus bacteriophages. *Scientific Reports.* 2019;9(1):7991.

98. **Ha AD, Denver DR.** Comparative Genomic Analysis of 130 Bacteriophages Infecting Bacteria in the Genus *Pseudomonas*. *Frontiers in Microbiology.* 2018;9(1456).

99. **Dion MB, Oechslin F, Moineau S.** Phage diversity, genomics and phylogeny. *Nature Reviews Microbiology.* 2020;18(3):125-38.

100. **Krumsiek J, Arnold R, Rattei T.** Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics.* 2007;23(8):1026-8.

101. **Moraru C, Varsani A, Kropinski AM.** VIRIDIC – a novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *bioRxiv*. 2020:2020.07.05.188268.

102. **Edgar RC.** MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics.* 2004;5:113-.

103. **Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF.** Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics.* 2011;12(1):395.

104. **Lamine JG, DeJong RJ, Nelesen SM.** PhamDB: a web-based application for building Phamerator databases. *Bioinformatics*. 2016;32(13):2026-8.

105. **Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, et al.** Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology.* 2019;37(6):632-9.

106. **Zhu W, Lomsadze A, Borodovsky M.** Ab initio gene identification in metagenomic sequences. *Nucleic Acids Research.* 2010;38(12):e132-e.

107. **Enright AJ, Van Dongen S, Ouzounis CA.** An efficient algorithm for large-scale detection of protein families. Nucleic acids research. 2002;30(7):1575-84.

108. **Buchfink B, Xie C, Huson DH.** Fast and sensitive protein alignment using DIAMOND. *Nature Methods.* 2015;12(1):59-60.

109. **Nepusz T, Yu H, Paccanaro A.** Detecting overlapping protein complexes in protein-protein interaction networks. *Nature Methods.* 2012;9(5):471-2.

110. **Hatfull GF.** Bacteriophage genomics. *Current opinion in microbiology.* 2008;11(5):447-53.

111. **Xie Y, Wahab L, Gill JJ.** Development and Validation of a Microtiter Plate-Based Assay for Determination of Bacteriophage Host Range and Virulence. *Viruses*. 2018;10(4):189.

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

113. **Duyvejonck H, Merabishvili M, Pirnay J-P, De Vos D, Verbeken G, Van Belleghem J, et al.** Development of a qPCR platform for quantification of the five bacteriophages within bacteriophage cocktail 2 (BFC2). *Scientific Reports.* 2019;9(1):13893.

114. **Anderson B, Rashid MH, Carter C, Pasternack G, Rajanna C, Revazishvili T, et al.** Enumeration of bacteriophage particles: Comparative analysis of the traditional plaque assay and real-time QPCR- and nanosight-based assays. *Bacteriophage*. 2011;1(2):86-93.

115. **Schmerer M, Molineux IJ, Bull JJ.** Synergy as a rationale for phage therapy using phage cocktails. *PeerJ*. 2014;2:e590-e.

116. **Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko C-C, et al.** Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. *J Mol Biol.* 2010;397(1):119-43.

117. **Roach DR, Leung CY, Henry M, Morello E, Singh D, Di Santo JP, et al.** Synergy between the Host Immune System and Bacteriophage Is Essential for Successful Phage Therapy against an Acute Respiratory Pathogen. *Cell Host & Microbe.* 2017;22(1):38-47.e4.

Section 7: Appendices

Appendix 1. Media Composition

Medium	Composition
Nutrient Broth	0.5% (w/v) Bacto yeast extract; 2.5% (w/v) Oxoid nutrient broth
Nutrient Agar	0.5% (w/v) Bacto yeast extract; 3.5% (w/v) Oxoid blood agar base
S.O.C.	0.5% (w/v) Yeast extract; 2.0% (w/v) Tryptone; 10mM MgSO4; 20mM D-Glucose

Appendix 2. Bacteriophage DNA extraction reagents

Reagent	Composition	Concentration
10% SDS	SDS	10% (w/v)
	Milli-Q water	90% (v/v)
70% Ethanol	Ethanol	70% (v/v)
	Milli-Q water	30% (v/v)

Appendix 3. Gel electrophoresis reagents

Buffer/reagent	Composition	Concentration
50 x TAE buffer	Tris	24.2 % (w/v)
	Glacial acetic acid	5.71% (w/v)
	Na2EDTA (pH 8.5)	3.72 (w/v)
1 x TAE buffer	50 x TAE buffer	1% (v/v)
	Milli-Q water	-
6 x DNA loading buffer	Tris-HCl	10 mM
	Bromophenol blue	0.03% (w/v)
	Xylene cynaol FF	0.03% (w/v)
	Glycerol	60% (v/v)
	EDTA	60 mM
GeneRuler 1kb DNA ladder	DNA	0.5 µg µl-1
	Tris-HCl (pH 7.6)	10 mM
	EDTA	1 mM
	6 x DNA loading buffer	80% (v/v)
λ DNA-HindIII ladder	DNA	0.5 µg µl-1
	Ficoll®-400	2.5%
	EDTA	10 mM
	Tris-HCl (pH 8.0)	3.3 mM
Agarose gel	Agarose	0.8% - 1 % (w/v)
	1 x TAE buffer	97.5% - 98.5% (v/v)
	SYBR-safe	1.5%

Appendix 4. iProof high fidelity PCR reagents

Buffer/ reagent	Components	Concentration
iProof buffer	Potassium chloride	2.5 – 5%
	Triton X-100	0.1 – 1%
	Water	50 – 100%
dNTPs mix	ATP	10 mM
	CTP	10 mM
	GTP	10 mM
	TTP	10 mM
DMSO	Dimethyl sulfoxide	10% v/v
iProof Polymerase	-	2 U µl-1