Genomic comparison of novel *Staphylococcus aureus* bacteriophage and their anti-biofilm properties against MRSA sequence type 22 and 36

Elliot Whittard

PhD 2020

Genomic comparison of novel *Staphylococcus aureus* bacteriophage and their anti-biofilm properties against MRSA sequence type 22 and 36

Elliot Whittard

A thesis submitted in partial fulfilment of the requirements of the Manchester Metropolitan University for the degree of Doctor of Philosophy

School of Healthcare Science

Manchester Metropolitan University

2020

Declaration and statements

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed………………………………………………(candidate)   Date………………

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed………………………………………………(candidate)   Date………………

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed………………………………………………(candidate)   Date………………

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for interlibrary loan, and for the title and summary to be made available to outside organisations.

Signed………………………………………………(candidate)   Date………………

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my director of studies Professor Mark Enright. It is hard to put in to words how grateful I am for the opportunity you have given me to pursue a PhD. You have been a constant source of inspiration and a great mentor, thank you for your continuous support and guidance throughout this whole journey.

My profound gratitude to Professor Jo Verran for your enthusiastic attitude towards my work, thank you for always being available to talk to whenever I needed some helpful advice.

I owe a great deal of thanks to Dr James Redfern, I am incredibly grateful for the infinite advice, encouragement and support that you have provided me on a daily basis. Your unabated enthusiasm and outstanding commitment to science has been inspiring and it has been an absolute pleasure to work alongside you.

I would also like to extend my gratitude to Dr Guoqing Xia and Dr Andrew Millard for sharing your suggestions and insights during my study. I must acknowledge the fellow academics, researchers and technicians at MMU who have made my time as a PhD student so much more enjoyable, especially Nat and Lesley for your helpful and somewhat questionable advice and for allowing me to pester you both every day. A massive thank you to John, Helen, Tony, Chris and Amanda for all the encouragement and support.

Thank you to my dear friends Mia and Ness, you two have been pivotal throughout my studies. There have been times where I have disappeared into the depths of my thesis writing, but you have always been there for me no matter what. I count myself incredibly lucky to have you as my friends.

A special thanks to Danielle for your incredible patience and understanding over the last couple of years, your endless support and encouragement has pushed me through this endeavour. I am truly grateful to have you in my life.

To my incredible parents and brother, your unwavering support throughout this whole journey has allowed me to achieve far more than I could have ever thought was possible. Thank you for believing in me and my abilities, even when I stopped believing in myself, I could not have completed this PhD thesis without you all.

Abstract

Staphylococcus aureus (including methicillin-resistant *S. aureus* - MRSA) remains a leading cause of both nosocomial and community acquired infections globally and despite constant improvement efforts to patient safety within healthcare, it still remains associated with considerable rates of morbidity and mortality. *S. aureus* is a common cause of biofilm-associated infections observed in chronic wounds, exhibiting a reduced susceptibility to the action of conventional antimicrobial agents and are often difficult to eradicate. The acquisition of resistance to almost any antibiotic with reference to MRSA has greatly reduced the number of alternative antimicrobial agents effective in the treatment of infections. Current development pipeline for new classes of antibiotics are greatly limited, requiring new, alternative approaches for therapeutic and prophylactic intervention in attempt to effectively control and overcome this current global health threat. Bacteriophage therapy exploits the natural killing ability of lytic bacteriophage (phage) as a means of controlling multidrug-resistant pathogenic bacteria. The utility of phage and their derivatives has been shown to effectively reduced the biofilms of major MRSA clones *in vitro* and *in vivo.* Global MRSA infections are caused by highly-successful isolates from a small number of epidemic lineages (clones). ST22 and ST36 are two of the most prevalent clones with global impact and largely responsible for the national epidemic of MRSA infections within UK healthcare system throughout the mid-1990s up until the mid-2000s. Understanding the phenotypic and genotypic characteristics of these clones in relation to the ability of bacteriophage to infect and disrupt established biofilms has yet to be explored.

In this study, a total of 46 novel obligately lytic phage were isolated from wastewater samples by utilising a modified *Staphylococcus carnosus* TM300 isolate with expressed *S. aureus* wall teichoic acids to aid in phage adsorption. The addition of 32 more phage from

our current laboratory stocks helped to establish a collection of 78 phage that were screened against a panel of 185 genetically diverse *S. aureus*, consisting of major clonal groups with high prevalence within the UK and United States, including 43 ST22 and 24 ST36 strains. The majority of the members displayed a wide host range against our panel. Based on this, the four most effective (wide host-range) phage were assessed for their anti-biofilm properties in polystyrene plates biofilm assays produced using four ST22 and four ST36 isolates. Treatment of mature biofilms was shown to significantly reduce biofilm biomass and viable cell counts. However these assays selected for the emergence of phage resistant mutants.

Whole genome sequencing was performed on 22 phage isolates and these were found to share a high degree of similarity to genomes of 38 previously classified *Twortvirinae* phages represented in GenBank. Comparisons of these 60 phage genomes found a surprisingly high level of genetic diversity. Pairwise distances resolved groups of phage in distinct clusters representing individual genera within the *Twortvirinae* subfamily. Pan-genome analysis identified no single gene present amongst all phage genomes, however phage displayed a core genome amongst other members of the cluster. The structural homology tool HHpred was used to predict the protein structure of genes encoding for lytic enzymes among our phage genomes. We found that all phage encode a protein that shares high structural similarity to the same CHAP domain protein, a catalytic domain of endolysins employed by phage to degrade the bacterial host cell wall in order thus, mediate cell lysis. Suggesting that the all phage most likely share the same catalytic N-terminal endopeptidase domain of endolysins which have a modular domain structure. Interestingly, endolysins have been proposed as possible candidates for the control of antibiotic resistant *S. aureus* infections.

Table of Contents

VIII

List of Figures

XI

List of Tables

Abbreviations and acronyms

Abi	Abortive infection
agr	Accessory Gene Regulation
AMR	Antimicrobial Resistance
ASP	Activated Sludge Process
BLAST	Basic Local Alignment Search Tool
bp	Base Pairs
BRIG	BLAST Ring Image Generator
CA-MRSA	Community-Acquired Methicillin-Resistant *Staphylococcus aureus*
CC	Clonal Complex
ccr	Cassette Chromosome Recombinase
CDS	Coding DNA Sequence
CFLM	Confocal Laser Scanning Microscopy
CFU	Colony Forming Unit
CHAP	Cysteine, Histidine-Dependant Amidohydrolase/Peptidase Domain
CONS	Coagulase-Negative *Staphylococcus aureus*
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPR-cas	CRISPR-associated proteins
CV	Crystal Violet Stain
DNA	Deoxyribonucleic Acid
Dnase	Deoxyribonuclease
dNTPs	Fluorescently-Labelled Deoxynucleotides Triphosphates
dsDNA	Double Stranded DNA
EMRSA-15	Epidemic MRSA-15, Also Known As CC22, ST22 Sccmeciv
EMRSA-16	Epidemic MRSA-16, Also Known As CC30, ST36 Sccmecii
EPS	Exopolysaccharide Substance
FnBP	Fibronectin-Binding Proteins
HA-MRSA	Hospital-Acquired Methicillin-Resistant *Staphylococcus aureus*
HGT	Horizontal Gene Transfer
ICTV	International Committee for Taxonomy of Viruses
IWG-SCC;	International Working Group On The Staphylococcal Cassette Chromosome Elements
kbp	Kilobase Pairs
Mbp	Megabase Pairs
MDR	Multi-Drug Resistance
MGE	Mobile Genetic Element
MIC	Minimum Inhibitory Concentration
MLST	Multi-Locus Sequence Type
MOI	Multiplicity of Infection
MRSA	Methicillin-Resistant *Staphylococcus aureus*
MSCRAMMS	Microbial Surface Components Recognizing Adhesive Matrix Molecules
MSSA	Methicillin-Susceptible *Staphylococcus aureus*

MUSCLE	Multiple Sequence Comparison by Log-Expectation
NCBI	National Centre for Biotechnology Information
NGS	Next Generation Sequencing
nm	Nanometer
nM	Nanomolar
OD	Optical Density
ORF	Open Reading Frame
PBP	Penicillin-Binding Protein
PBS	Phosphate Buffered Saline
PDP	Protein Data Bank
PFU	Plaque Forming Unit
PIA	Polysaccharide Intercellular Adhesion
PVL	Panton-Valentin Leukocidin
RboP	Polyribitol-Phosphate
R-M system	Restriction Modification system
RNase	Ribonuclease
Rpf	Resuscitation Promoting Factor
rpm	Rotations per Minute
S. aureus	*Staphylococcus aureus*
S. carnosus	*Staphylococcus carnosus*
S. epidermidis	*Staphylococcus epidermidis*
SAgs	Staphylococcal Superantigens
SaPI	Staphylococcal Pathogenicity Island
SCC	Staphylococcal Cassette Chromosome
SCCmec	Staphylococcal Cassette Chromosome Carrying the *mec* Gene
SD	Standard Deviation
SRA	Sequence Read Archive
SSSS	Staphylococcal Scalded Skin Syndrome
ST	Sequence Type
TSA	Tryptic Soy Agar
TSB	Tryptic Soy Broth
TSBg	TSB Supplemented with 1 % D-(+)-Glucose
TSST-1	Toxic Shock Syndrome Toxin-1
UK	United Kingdom
v/v	Volume For Volume
VISA	Vancomycin-Intermediate *Staphylococcus aureus*
w/v	Weight per Volume
WGS	Whole Genome Sequencing
WTA	Wall Teichoic Acid
WwTW	Wastewater Treatment Works
µg	Micrograms
µl	Microlitres
µm	Micrometre

Chapter 1

Literature review

1.1 Current situation

The development and introduction of antibiotics to treat bacterial infections was one of the key medical advances of the 20th century. Society relies heavily upon safe and effective antibiotics to treat an array of infections, but unregulated use, overuse and misuse within human healthcare, veterinary medicine and in food production has led to the evolution and spread of antibiotic resistance within all human pathogenic bacterial species [1]. Resistance to first choice, frontline antibiotics leads to delays in effective treatment and alternative drugs, if available may be more toxic, less effective and costlier and result in increased morbidity, mortality and increased hospital stays. Based on data acquired in 2015, researchers estimate that more than 33,000 deaths per year in Europe are attributable to healthcare-associated multidrug-resistant (MDR) infections due to pathogens that were once curable but are becoming increasingly difficult to treat [2]. Using current rising rates, the number of deaths attributed to drug-resistant pathogens is predicted to balloon to 10 million deaths per year presenting a major, growing public health challenge requiring a global response (Figure 1.1) [3].

There are over 15 classes of antibiotics that target different bacterial physiological and metabolic processes resulting in cell death or cessation of growth. Resistance mechanisms have evolved and spread to all these drug classes, with recent reports of strains even exhibiting resistance to 'last resort' antibiotics such as vancomycin, carbapenems and polymyxins [3].

Many pathogens are resistant to more than one antibiotic class and alternative agents may be prohibitively expensive especially in developing countries. These alternative antibiotics may be more toxic, more poorly absorbed and less effective than frontline therapies [4].
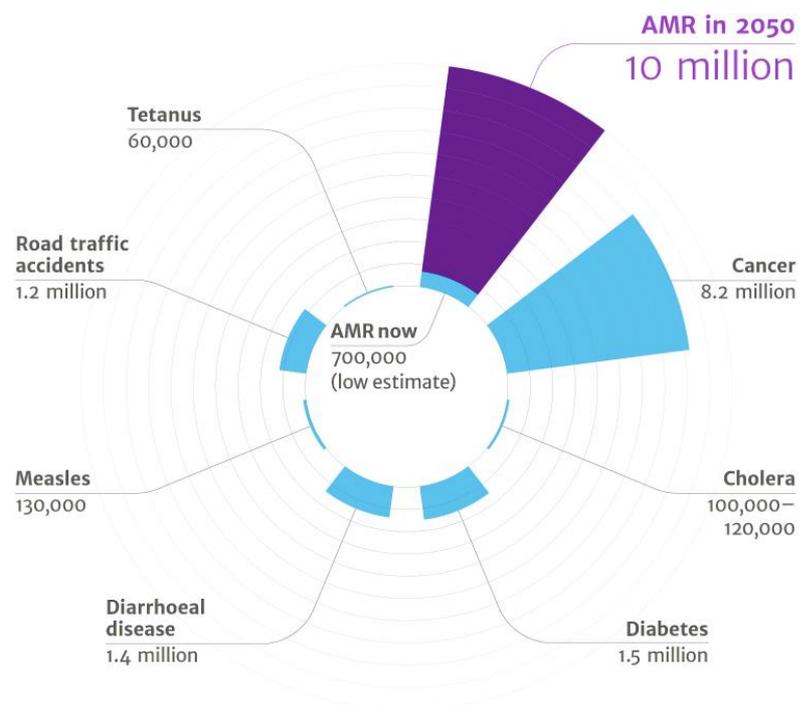
Figure 1.1: Predicted number of deaths worldwide attributable to antimicrobial-resistant pathogens each year by 2050. Figure adapted from [384].

The current problems caused by the evolution and spread of AMR pathogens has dramatically increased the need for safe and effective alternatives adjuncts to conventional antimicrobial chemotherapies [5]. Several alternatives to antibiotics do exist but are not yet widely available, these include bacteriophage therapy, immunotherapy and antimicrobial peptides. Research and development efforts in

these areas offers the possibility of regaining control over AMR infections and prolonging the utility of antibiotics [4].

1.2 *Staphylococcus aureus*

Staphylococcus aureus remains as one of the primary causes of both hospital and community-acquired infections globally, an extremely versatile pathogen, it is responsible for causing a broad spectrum of diseases ranging in severity from minor skin and soft-tissue infections such as boils, furuncles and impetigo to more life-threatening invasive infections including pneumonia, osteomyelitis, endocarditis, meningitis and septicaemia [6]. It's versatility has enabled the pathogen to quickly develop and acquire resistance to all antibiotic classes, whilst demonstrating an array of virulence factors that facilitates host colonisation and immune evasion such as cytolytic exotoxins, adhesins and biofilms production [7].

S. aureus is a frequent coloniser of the skin and mucous membranes of mammals with the anterior nares being its primary niche in humans. The epidemiology of *S. aureus* infections is multifactorial with both bacterial virulence factors and host factors thought to play a significant role. Around 30 % of the population are estimated to be asymptomatically colonised with this pathogen which may persist for long periods of time [8,9], with a further 60 − 80 % being intermittently colonised [10]. Despite being a normal component of the host microflora, it is considered an opportunistic pathogen as once it contaminates a breach in the skin or mucous membranes of the host, it can take advantage of the impaired host immune response and deploy a repertoire of host pathogenic molecules such as

hyaluronidase and exotoxins to break down host tissues, enabling the pathogen to invade deeper tissues [11]. Hospitalised patients are particularly prone to *S. aureus* infections due to the presence of compromised immune systems, surgical site infections caused by the implantation of indwelling medical devices such as protheses, stents and venous catheters [12]. Often leading to extended hospitalisation periods with increased risk for subsequent disease development, whilst causing significant financial burden to healthcare services [13].

1.2.1 Pathogenicity

The success of *S. aureus* as a major human pathogen is largely due to its extensive array of virulence determinants present among strains that are used to thwart host immune responses. The majority of which are encoded on mobile genetic elements (MGEs) such as pathogenicity islands (PIs), plasmids, bacteriophages and transposons that are an important means for transfer of genetic information within and between species [14]. MGEs play a significant role in the genome plasticity not only contributing to diversity among individual strains, but also allows the bacteria to adapt to constant changes in environment and its human host [15]. Horizontal gene transfer of virulence factors encoded within MGEs between strains has resulted in multiple combinations of virulence factors, that associate with particular host genetic backgrounds and clinically significant clones [16]. These virulence factors can be classed into three functional groups [17]:

- Surface proteins that mediate adhesion of bacteria to cells or tissues.
- Extracellular enzymes which promote tissue damage and spread.

- Toxins to help protect the bacteria from the host immune system.

S. aureus possess the ability to produce a comprehensive repertoire of proteins that are collectively known as microbial surface components recognizing adhesive matrix molecules (MSCRAMMS) [18]. These surface proteins aid and promote the adherence and attachment to the surface of cells, tissues and prosthetic devices. MSCRAMMS are secreted primarily during exponential growth contributing to the pathogenesis of *S. aureus* and helping to establish infection [11,19]. Another protein known to mediate immune evasion is the surface associated protein A (spA), an immunoglobulin G (IgG)-binding protein present on the cell wall, which binds to the Fc regions of IgG subsequently preventing opsonophagocytosis (phagocytosis of bacteria or foreign cell initiated by an antibody) [20]. Other examples of MSCRAMMs include the extracellular adherence protein (EAP) which have been shown to promote invasion of eukaryotic cells and also bind to and interact with numerous plasma proteins including fibrinogen and fibronectin [21]. Two closely related fibronectin-binding proteins (FnBPs) FnBPA and FnBPB are associated with inflammation and sepsis as well as facilitate the internalisation of *S. aureus* to host cells and attachment of the bacterial cell to fibronectin, elastin and fibrinogen [19,22].

1.2.2 Toxins

Most *S. aureus* isolates are capable of producing an array of extracellular enzymes such as coagulase, proteases, haemolysins, hyaluronidase and collagenases [7]. These enzymes contribute to the spread and pathogenesis of bacteria against host cells. Additionally, *S. aureus* produce and secrete several cytotoxins in post-exponential and early stationary phase. These are involved in tissue penetration and induce pro-inflammatory responses that facilitates the spread and dissemination of infection specifically targeting human cells and are grouped according to their mode of action as depicted in Figure 1.2 [23].

EXPONENTIAL PHASE

CELL SURFACE VIRULENCE FACTORS
1. Protein A
2. Fibronectin-binding protein
3. Clumping factor
4. Other MSCRAMMs

STATIONARY PHASE

SECRETED VIRULENCE FACTORS
1. Superantigens (SEs, SEIs, TSST-1)
2. Cytolysins (α, β, γ, δ-toxins, PVL)
3. Exoenzymes
4. Polysaccharide intercellular adhesin (slime)

Staphylococcus aureus

CELL WALL
1. Peptidoglycan
2. Lipoteichoic Acids, Teichoic Acids
CAPSULE

Figure 1.2: The virulence factors produced by *Staphylococcus aureus*, figure adapted from [385]

Hemolysins (α, β,γ and δ) are porin-like toxins that lyse various host cells including erythrocytes, platelets and monocytes. There are four types of hemolysins which are responsible for diseases such as food poisoning, toxic shock syndrome and exfoliative shock syndrome toxin-1 (TSST-1). Probably the best characterised and

studied cytotoxin produced by *S. aureus* is α-toxin, which is capable of binding to target cell membranes and forming pores in the lipid bilayer subsequently resulting in cell lysis [7,24].

Leukotoxins target white blood cells, forming pores in the polymorphonuclear cell membranes of neutrophils and macrophages [25]. Panton-Valentine leukocidin (PVL) is probably the most clinically significant toxin produced by *S. aureus* responsible for causing necrotising pneumonia, leukopenia and cutaneous infections [26–28]. The PVL toxin is encoded by a prophage incorporated into the *S. aureus* chromosome (see section 1.2.5.2), PVL positive isolates originate from genetically diverse backgrounds and are associated with outbreaks of community-associated MRSA infections [29,30].

Exfoliative toxins (ETs), also referred to as epidermolytic toxins are serine proteases that hydrolyse desmosome proteins present only in the superficial layers of the skin and are primarily responsible for staphylococcal scalded skin syndrome (SSSS). A generalised exfoliative disease affecting infants and children, SSSS manifests with fever and malaise, leading to blistering and severe exfoliation of skin surface [31].

Staphylococcal superantigens (SAgs) are a family of exotoxins that induce T-cell proliferation resulting in uncontrolled release of cytokines that can lead to severe tissue inflammation, hypovolemic shock, multiple organ failure and death [32]. These toxins include the notable toxic-shock syndrome toxin-1 (TSST-1) and

staphylococcal enterotoxin C (SEC) that play a critical role in the aetiology of *S. aureus* infective endocarditis [33]. Genes encoding for these superantigens are usually found within *S. aureus* pathogenicity islands (SaPI) as detailed in section 1.2.5.1.

1.2.3 Biofilms

Bacteria seldom exist in isolation and are more commonly found inhabiting various environments as mixed microbial communities consisting of more than one bacterial species termed as polymicrobial [34,35]. Biofilms are commonly defined as an assemblage of sessile microorganisms enclosed within a complex exopolysaccharide substance (EPS) or glycocalyx layer that accounts for up to 90 % of the biofilm biomass and acts as a cell survival mechanism [36]. The production of biofilms is an effective survival strategy employed by almost all bacterial species and has been extensively studied as these cells are often difficult to treat, largely owing to the protective nature of the glycocalyx layer that antibacterial agents have trouble penetrating and reaching its target [37]. Cells embedded deep within the biofilms layers tend to demonstrate a different physiology and biochemistry in comparison to their planktonic counterparts [38]. Such changes can be observed with the presence of metabolically inactive persister cells that exhibit reduced growth rates. This greatly reduces the uptake of antimicrobials that are designed to target fast growing and metabolically active bacteria, thus further contributes to their survival [39,40]. Biofilms cells are substantially more difficult to treat with conventional antibiotics and sanitising agents, often withstanding concentrations up to 1,000 times greater than those tolerated by planktonic cells [41]. The

formation of biofilms is classically described as a multistage process consisting of four key phases: adherence, accumulation, maturation and dispersal (Figure 1.3). There are a number of variables that mediate the initial attachment of bacteria to a surface and is a critical step in establishing an infection.
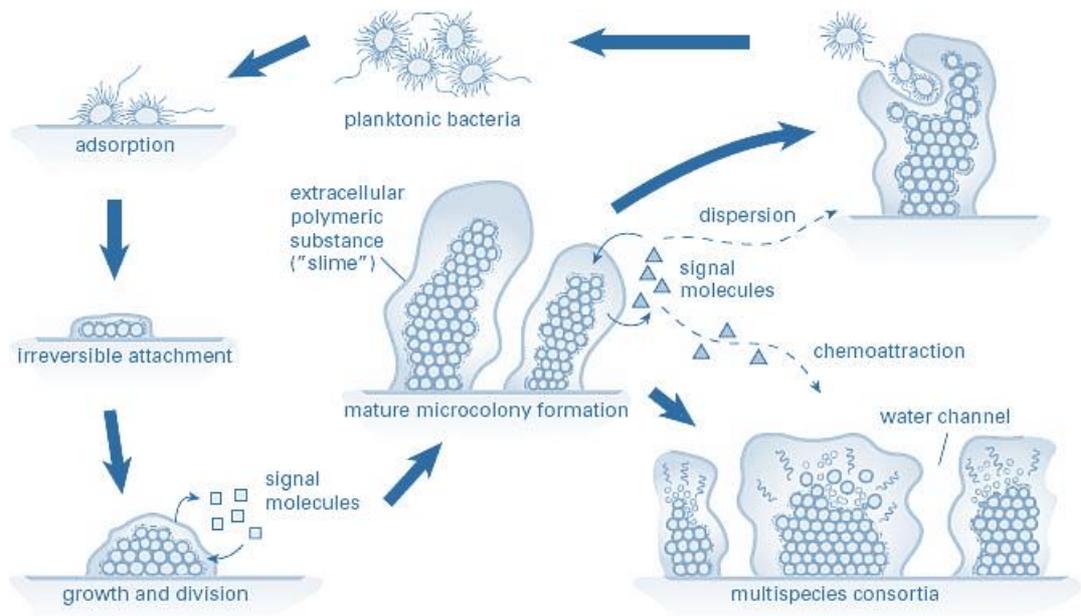
Figure 1.3: The multiple stages of biofilm formation and development, figure adapted from [386].

1.2.3.1 Adhesion

Adherence between *S. aureus* and abiotic surfaces such as indwelling medical devices is mediated by non-specific interactions such as hydrophobic, Van der Waals and electrostatic forces [42]. Whereas, adhesion with human tissue is generally achieved by utilising one of the many adhesins encoded by *S. aureus* [43]. Adhesion and initial biofilm formation is achieved using a number of the proteins covalently anchored to the cell surface such as the MSCRAMMs, non-covalently associated proteins such as autolysins and non-proteinaceous surface-associated

proteins that include wall teichoic acids (WTA) and polysaccharide intercellular adhesion (PIA) to recognise and interact with the host extracellular matrix components [18,43]. The surface is rapidly coated in host matrix proteins such fibrinogen, fibronectin and collagen which further encourages cell adhesion and colonisation [44]. Once attached a change in gene expression is observed allowing *S. aureus* cells to initiate irreversible attachment mechanisms [45].

1.2.3.2 Accumulation

Following initial attachment, cells start to multiply and accumulate together in to small multilayered aggregates called microcolonies. As cells begin to proliferate, they are linked together by adhesive biofilm molecules known as polysaccharide intercellular adhesin (PIA), also termed as poly-N-acetyl-glucosamine (PNAG) and are thought to play a crucial role in maintaining the stability of the immature biofilms [46]. In the absence of EPS, newly formed microcolonies are vulnerable to detachment particularly in the presence of shear forces such as flowing fluid [47]. PIA is partially de-acetylated resulting in a net positive charge, thus attracting the negatively charged bacterial cell surface partly provided by the WTA and subsequently adhering cells together through electrostatic interaction [48]. Cells begin to produce biofilm extracellular components forming a three-dimensional complex encasing the microcolony [45].

1.2.3.3 Maturation

Biofilms develop as the cells continue to replicate, the agglomeration of microcolonies form in to large multi-layered biofilm structures increasing in density

and complexity. The glycocalyx layer of the biofilm begins to develop from the production of extracellular components by the adhered cells interacting with the organic and inorganic molecules within the immediate environment [49]. Bacteria use quorum sensing to regulate expression of genes that are actively involved and required during biofilm development. Quorum sensing is a cell-cell signaling mechanism that regulates gene expression in a cell-density dependent manner. It is mediated by hormone-like diffusible molecules called autoinducing peptides (AIP), whereby upon accumulation are recognized by cell surface receptors. The structured organization of cells allows the formation of interstitial voids and distinct water channels devoid of bacterial cells within the complex EPS structure separates microcolonies, these channels are necessary to allow the supply of nutrients, gas diffusion, signaling molecules to reach cells in the deeper layers of the biofilm as well as the removal of waste products [49,50]. However a number of factors including nutrient availability, internal pH and oxygen perfusion continue to limit the growth potential of the biofilm [49]. Once fully developed, the biofilm exopolysaccharide matrix is composed primarily of teichoic acids, polysaccharide intercellular antigen (PIA), extracellular DNA (eDNA) and staphylococcal/host proteins [47].

1.2.3.4 Dispersal

The detachment and dispersal of mature, fully developed biofilms have been found to be an important step in the dissemination of infection in attempt to establish secondary biofilm infections elsewhere within the host or possibly non-biofilm associated systemic infections such as septicemia and infective endocarditis [51].

Dispersal of biofilm cells has been found to be mediated by the accessory gene regulator (*agr*) locus of the quorum sensing system which controls virulence within *S. aureus* [52]. Activation of the *agr* system triggers a regulatory cascade suppressing the development of biofilms by downregulating the expression of surface adhesion proteins and by upregulating the expression of extracellular degrative enzymes such as proteases which facilitate detachment of biofilms, reverting cells back to a planktonic state thus encouraging the dispersal [48,52].

1.2.4 *S. aureus* biofilms

S. aureus is one of the most common causes of biofilm-associated and device-related infections within healthcare, the bacterium presents a burden for healthcare systems and patient safety as it commonly colonises chronic wounds and the surfaces of indwelling medical devices such as intravenous catheters, cardiac pacemakers and joint prostheses [53,54], leading to biofilm-related diseases such as osteomyelitis, periodontitis and endocarditis [43]. Such infections are often difficult to treat using conventional antibiotics, with medical devices having to be replaced much more frequently compared to those infected by other staphylococcal species [55]. A notable characteristic of *S. aureus* biofilms is their ability to produce vast amounts of EPS, composed of teichoic acids, polysaccharide intercellular antigens (PIA), DNA and staphylococcal / host protein components [29]. Their success in biofilm production can be linked to their ability to produce an array of proteins known as MSCRAMMs that can help manipulate the human host defence, aid in adherence and attachment to the surface of cells, tissues and

prosthetic devices [18]. Their decreased susceptibility to antimicrobial agents contributes to their persistence in infections.

S. aureus biofilms can be a major foci of infection with detachment and dispersal of aggregates through the vascular system leading to the colonisation of new niches. This can result in progression of disease from, for example, wound infection to disseminated / systemic disease including endocarditis, necrotising pneumonia, meningitis and septicaemia [56]. *S. aureus* is endemic in hospital environments, and is also a major cause of community-associated infections, these are typically skin and soft tissue MSSA infections but in some locales and population groups community-associated MRSA infections (CA-MRSA) are common in the absence of the risk factors associated with healthcare-associated infections [57].

1.2.5 Evolution and genetic diversity

The genomes of *S. aureus* are circular chromosomes that are on average around 2.8 Mbp in length, however the length can vary depending on the size of additional extrachromosomal plasmids. Comparisons among *S. aureus* genomes typically reveals a core genome of roughly 78 % shared among isolates. The core genome refers to genes that are present across all strains, these orthologous genes are generally highly conserved and are essential for cellular metabolism, replication, surface-expressed proteins and genes encoding housekeeping functions [58]. Additionally, core-variable genes, a group of relatively stable genes that are variably shared between isolates and likely transferred vertically make up approximately 10 − 12 % of any *S. aureus* genome [58]. These core-variable genes

are mainly putative virulence or resistance genes carried on mobile genetic elements (MGEs), several unique combinations of such genes were found overrepresented in a number of specific clonal lineages inhabiting different ecological niches. Suggesting evolutionary and functional adaptation correlating with their pathogenicity and prevalence [59–61].

The accessory components of the genome account for approximately 25 % of the total genome and is where the majority of genetic variation within a bacterial species is observed. Primarily consisting of MGEs including bacteriophage, staphylococcal pathogenicity islands (SaPIs), plasmids, transposons and staphylococcal cassette chromosomes (SCC), these components are frequently exchanged through horizontal gene transfer (HGT) despite the highly clonal population of *S. aureus*. The gain and loss of MGEs facilitates the rapid adaptation of bacteria to an ecological niche and are an important source of plasticity and diversity of the genome by driving bacterial evolution. Selective pressures from the environment drives the enrichment and need for specialised genes that promote fitness including virulence determinants or aid in the survival of the bacteria, such as acquiring genes that confer resistance to antimicrobial drugs [62].

1.2.5.1 S. aureus pathogenicity islands (SaPIs)

The majority of *S. aureus* pathogenicity islands (SaPIs) encode for one or more staphylococcal superantigens (SAgs) and are primarily responsible for causing SAg-induced diseases such as TSST-1. These chromosomal elements range from 14 −

17 kilobase pairs in size and reside at one of six specific att sites on the chromosome [63,64]. SaPIs can be mobilised and disseminated amongst populations by the use of certain helper phage or induction by prophage and are an effecting method for genetic variation and transfer playing a significant role in the pathogenesis and virulence between strains [15].

1.2.5.2 Bacteriophage

Bacteriophage integrated within the bacterial chromosome (prophage) are widespread within *S. aureus* genomes, with almost all strains carrying at least one phage, with some possessing up to four in some cases. The prevalence and distribution of prophage has been linked to dominant *S. aureus* clonal complexes [65]. Phage encode a large portion of the virulence factors possessed by *S. aureus* strains such as PVL (encoded by the PV-*luk* operon), staphylokinase (*sak*) and enterotoxin A (*sea*) [66]. Phage are the primary vehicles for lateral transfer of genes among strains following activation and mediation between hosts, spreading chromosome-encoded virulence factors through generalised or specialised transduction. They play a significant impact on staphylococcal diversity and are tightly linked to host evolution and emergence of new pathogenic clones such as the community-associated MRSA strains.

1.2.5.3 Staphylococcal cassette chromosomes (SCC)

Staphylococcal cassette chromosomes (SCC), such as SCC*mec* are a defining feature of methicillin-resistant *S. aureus* (MRSA) and not found present within methicillin-susceptible *S. aureus* (MSSA) isolates. These mobile genetic elements

vary in size from 21 – 67 kb, and there are currently 11 (I – XI) different SCC*mec*

variants that have been described by the International Working Group on the

staphylococcal Cassette Chromosome elements (IWG-SCC;

http://www.sccmec.org/index.php/sccmec-uptodate - accessed 12-05-2019).

Each of which grant various phenotypic characteristics such as antimicrobial

resistance and are distinguished by the cassette chromosome recombinase (*ccrA*,

ccrB and *ccrC*) gene complexes. There are eight various types of *ccr* gene

complexes within staphylococci (1 - 8) and are responsible for the horizontal

transfer and dissemination of the elements within the chromosome locus,

contributing to the evolution of the SCC*mec* types and regulatory background

*mec*A architecture. There are currently six types of *mecA* gene complexes (class A,

B, C1, C2, D and E) with each gene composed of *mec*A and its regulatory genes

mecR1 and *mecI*.

1.3 Antibiotic Resistance

Resistance to methicillin is mediated by the penicillin binding protein (PBP2A), this

protein confers a lower affinity to all β-lactam antibiotics and as a result,

preventing the disruption of the peptidoglycan layer and synthesis, thus allowing

the bacteria to continue growing when in the presence of the antibiotic. PBP2A is

encoded by the mecA gene which resides on the SSC*mec* element, following the

acquisition of an SSC*mec* element, MRSA undergoes several mutational events

allowing the bacteria to become one of the most difficult to treat pathogens with

conventional antibiotics.

1.3.1 Hospital-acquired methicillin-resistant *S. aureus* (HA-MRSA)

The versatility of *S. aureus* has enabled it to withstand and adapt to the selective pressures of antimicrobial drug classes, by the start of the 1960s, roughly 80 % of *S. aureus* isolates exhibited resistance to penicillin. Resistance caused by the acquisition of a plasmid-encoding penicillinase led to the development of a semi-synthetic β-lactam antibiotic known as methicillin (or meticillin). Introduced in to clinical practice in 1959 to treat infections caused by penicillin-resistant *S. aureus*, however by 1961, the first reports of nosocomial methicillin-resistant *S. aureus* (MRSA) isolates were already starting to be described [67]. Since then MRSA clones have subsequently emerged and disseminated worldwide becoming a global concern and still remains accountable for considerable rates of nosocomial infections [68].

SCC*mec* elements I - III are usually found within hospital-acquired MRSA (HA-MRSA) isolates, possessing multiple combinations of ccr and *mec*A gene complexes that mediate antibiotic resistance [69]. Although predominantly restricted to healthcare, other successful MRSA clones possessing variants of SCC*mec* began to arise throughout the community in the late 1990s.

1.3.2 Community-associated MRSA

S. aureus is also a major cause of community-associated infections, such infections typically include minor skin and soft tissue diseases up to more severe and life-threatening conditions such as, septicaemia, necrotising pneumonia [70,71]. The

prominence of MRSA lineages not related to HA-MRSA are appearing within the community at an increasing rate, and are frequently transmitted among individuals with no predisposing conditions or previous contact within a healthcare setting, often making it difficult in establishing the origin of the MRSA isolate. By contrast, CA-MRSA isolates can be distinguished from their nosocomial counterparts as they often carry smaller SCC variants such as SCC*mec* type IV and V, the smaller elements are more mobile and may be advantageous when transferring between other bacteria. Unlike HA-MRSA, CA-MRSA isolates tend to be more susceptible to non- β-lactam drug classes and are more likely to carry extra virulence determinants such as Panton-Valentine Leukocidin (PVL) encoding genes [72]. CA-MRSA is becoming increasingly widespread across many countries around the world including the UK and USA, most notably the dominant epidemic clones USA 300 (ST8-IV) and European CA-MRSA (ST80-IV) [73]. Furthermore, CA-MRSA has been spreading in epidemic proportions in recent years, leading to an increase in CA-MRSA incidences within hospitals being reported [68]. Not only does this further complicate the procedures in place for infection control within public health, but also suggests evidence of clonal spread and crossover between HA-MRSA and CA-MRSA strains [74,75]

1.3.3 Global distribution of MRSA populations

1.3.3.1 Characterisation of MRSA

Molecular typing plays an essential role in understanding the evolutionary and genetic relatedness of *S. aureus* populations. For MRSA it is an invaluable tool for

epidemiology studies allowing transmission cases to be traced and outbreaks related to specific clones to allow for preventative measures to be implemented.

Multi-locus sequence typing (MLST) is one of the most widely used typing techniques employed to study the evolutionary history of *S. aureus* and MRSA isolates and their variation over time. MLST is a highly discriminatory and reliable approach used to characterise *S. aureus* isolates based on the internal fragments of seven housekeeping loci (*arcC, aroE, glpF,gmk, pta, tpi* and *ygi*) ranging from 450 – 500 base pairs (bp) [76]. This sequence based approach analyses polymorphic variation of each gene fragment and assigns it to an allele number. The combination of these numbers for each housekeeping gene generates a numerical allelic profile (Table 1.1), which can then be used to determine its sequence type (ST) [76]. When two isolates are found to share identical sequences for these seven housekeeping genes and assigned in to the same ST, they are considered a clone.

However, if closely related isolates that share the same allelic profile for five or six of the seven housekeeping loci, these are classified as members of the same clonal complex (CC) [77]. To date, the MLST database houses roughly 3110 unique *S. aureus* sequence types (http://saureus.beta.mlst.net/, accessed on 27th May 2019).

Table 1.1: *Allelic profile of seven housekeeping loci used to determine the sequence type (ST) of S. aureus isolates employed by multi-locus sequence typing (MLST).*

arcC	*arcC*	*glpF*	*gmk*	*pta*	*tpi*	*ygi*	Sequence Type (ST)	Clonal Complex (CC)
1	4	11	4	12	1	10	ST5	CC5
3	3	1	1	4	4	3	ST8	CC8
7	6	1	5	8	8	6	ST22	CC22
2	2	2	2	3	3	2	ST36	CC30
10	14	8	6	10	3	2	ST45	CC45

SCC*mec* typing technique is used to detect regions of the SCC*mec* variant specific for that MRSA isolate, this approach identifies the unique combination of *ccr* and *mec* gene complexes used to classify the 11 SCC*mec* variants [78]. SCC*mec* typing has been employed in concert with MLST and has become an important tool used to discriminate between closely related MRSA and MSSA clones, including between HA-MRSA and CA-MRSA. Comparisons between the genomes of MRSA isolates from various geographical locations reveal a stable component present in all clones derived from a small and genetically distinct group [77–79]. The nomenclatural system currently employed for MRSA is based on MLST profiling in conjunction with SCC*mec* type; for example the EMRSA-16 clone MRSA252 is (ST36-II) [80].

1.3.3.2 Global distribution of MRSA populations

Molecular typing techniques have revealed that the population structure of *S. aureus* is highly clonal, with a relatively stable core genome, diversity is mainly due

to the accumulation of single nucleotide polymorphisms [77]. The rapid emergence of *S. aureus* infections and MRSA incidences in many parts of the world has partly been due to the dissemination of global clones. In fact, the majority of dominant isolates responsible for epidemic outbreaks of HA-MRSA (EMRSA) were associated with five distantly related but successful clonal complexes (CCs) CC5, CC8, CC22, CC30, and CC45 [81,82]. Which can also be described by their founder sequence type (ST) and descendants of pre-existing clones, or originating from epidemic MSSA lineages which had acquired the SCC*mec* element (Table 1.1) [80]. These clonal complexes can emerge when prevalence rates of independently evolving lineages rise within the population, which in turn diversifies through relative contributions of random genetic drift and recombination generating closely related variants that when combined form a clonal complex [77].

The distribution of major clones varies geographically and their dominance shifts over time, eventually replaced by another clonal group through epidemic waves that correspond to evolutionary changes in the SCC*mec* within a clonal complex, which further supports the theory that horizontal transfer plays a key role in the dissemination of methicillin resistance [80,83]. The proportion of nosocomial HA-MRSA incidences can vary considerably in each country, with low prevalence rates across northern European countries ranging from less than 1 % in Sweden, up to 53 % in Romania and Portugal [84,85]. Further afield, in countries where preventative control measures are much lower can see higher prevalence of infection rates of more than 70 % in Asian countries such as Vietnam, Korea and China, even up to 87 % such as in Sri Lanka in some cases [75].

1.3.4 Methicillin-resistant *Staphylococcus aureus* (MRSA) in the United Kingdom

In the UK, while a number of notable MRSA clones have circulated, the MRSA epidemic dominating within hospitals and the community during the mid-1990s up until the mid-2000s can be largely associated with two waves of epidemic clones, UK EMRSA-15 (ST22-IV/CC22) and UK EMRSA-16 (ST36-II/CC30), both of which are collectively accountable for roughly 95 % of invasive hospital- and community-acquired MRSA bacteraemia reports throughout that time [80,86]. Their success at surviving and spreading throughout the UK can be largely due to the distinctive core genome and possession of unique accessory genes that contain several staphylococcal enterotoxin genes including seg and sei [87].

1.3.4.1 Sequence Type 36 (EMRSA-16)

Members of the ST36 (CC30) lineage include the pandemic clone EMRSA-16, which is responsible for many major outbreaks circulating around hospitals within the UK [88]. The first observations of EMRSA-16 in 1992 were reported within three hospitals in the town of Kettering in Northamptonshire [89], however it has been estimated that the most recent common ancestor to EMRSA-16 may have emerged roughly 22 years earlier around 1975 [90]. The rapid dissemination of EMRSA-16 around the UK was observed during a survey in 1995 (Figure 1.4), revealing 443 outbreaks of infection within hospitals caused by the clone occurred over 18 months, with surgical, medical and elderly-care wards accounting for 42 % (14 % each) of those reports. In the same study, patients colonised or infected with EMRSA-16 were found among 20 % of 150 nursing homes within the community.

Over the following 3 years, EMRSA-16 became widespread in most regions of the UK accounting for 25 - 30 % of all MRSA isolates across the south of England, 23 % in the Midlands and 12 % in the north, as well as 25 % and 30 % in Scotland and Wales, respectively [89]. Bayesian phylogenetic reconstruction using a number of UK CC30 isolates resolved them into three clades representing pandemic clones southwest pacific (SWP), phage type 80/81 and EMRSA-16 originating independently instead of sequentially [90]. Revealing several subclades of EMRSA-16 (ST36) isolates unique to regions of the UK, correlating with the existence of
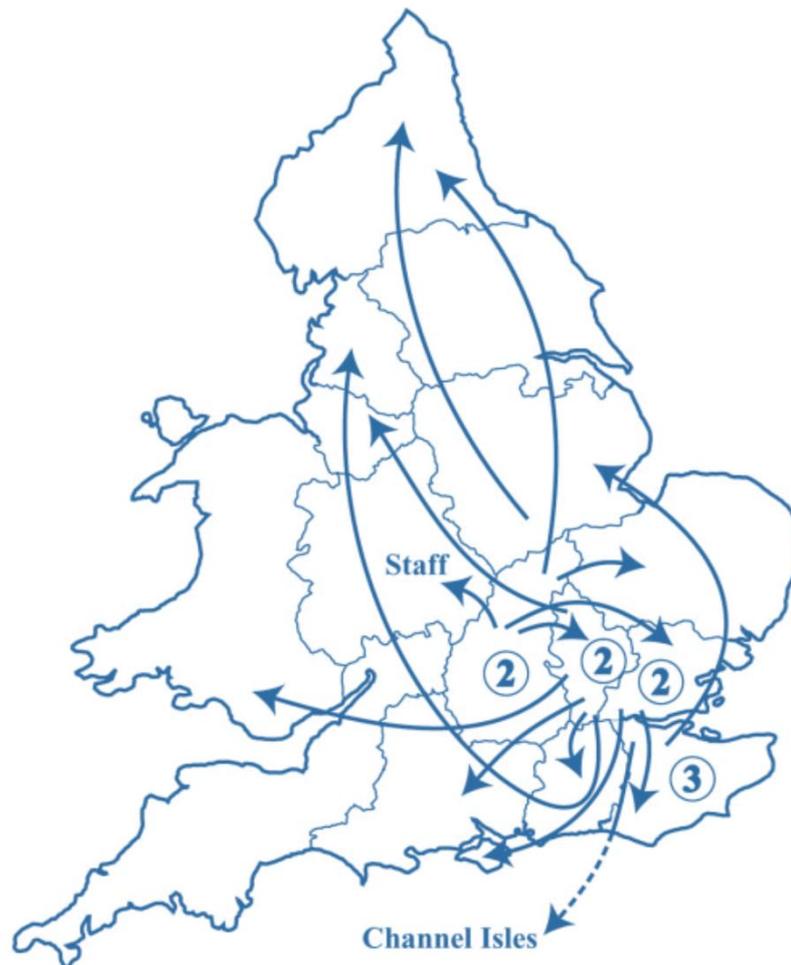
Figure 1.4: The initial reported route and rapid dissemination of EMRSA-16 colonisation by 24 of 136 hospitals from 1992 – 1994 following its emergence in Kettering, Northamptonshire. Figure adapted from [89].

EMRSA-16 isolates endemic to a particular hospital or region [90]. EMRSA-16 was commonly isolated in the UK, throughout the 1990s, however its prevalence within the UK rapidly declined from 21 % to 9 % between 2001 and 2007 and is not perceived as successful as the contending ST22 clones within the UK. Reports suggest that changes to infection control measures and antibiotic prescribing were possible factors for the decline of EMRSA-16 [91]. A shift in clonal structure of MRSA population across the UK was observed after 2003 as the dominant ST22 clone subsequently took over ST36 [92].

1.3.4.2 Sequence Type 22 (EMRSA-15)

ST22-IV-MRSA also known as the epidemic MRSA 15 (EMRSA-15) clone became the most successful and dominant MRSA clone associated with invasive disease in healthcare within the UK. Following the first reports in the early 1990s [93], EMRSA-15 outbreaks were reported across several European countries including Germany, Portugal and Czech Republic [94–96]. It has since then disseminated across to Australia, the Middle- and Far- East [91], where it has been found to replace previously dominant HA-MRSA lineages including CC5 and ST239 [97–101]. The reason for EMRSA-15 (ST22) and its dominance is not entirely understood, however it has been speculated that the smaller more mobile SCC*mec*IV element carried by the clone compared to the larger, more burdensome SCC*mec*II element in ST36 is thought to play a key role behind the success of EMRSA-15 [92]. This flexibility may allow the rapid adaptation to the selective pressures of multiply antibiotics and their various targets increasing its ability to acquire an array of new resistances such as ciprofloxacin (fluoroquinolone), which was introduced in to

clinical medicine within the UK in 1987 [92], clinal studies of the antibiotic were performed in the midlands prior to licensing. It might be argued that the EMRSA-15 isolates are generally biologically fitter due to the smaller MGE without carrying large resistance genes that often come at a cost to fitness and growth rate.
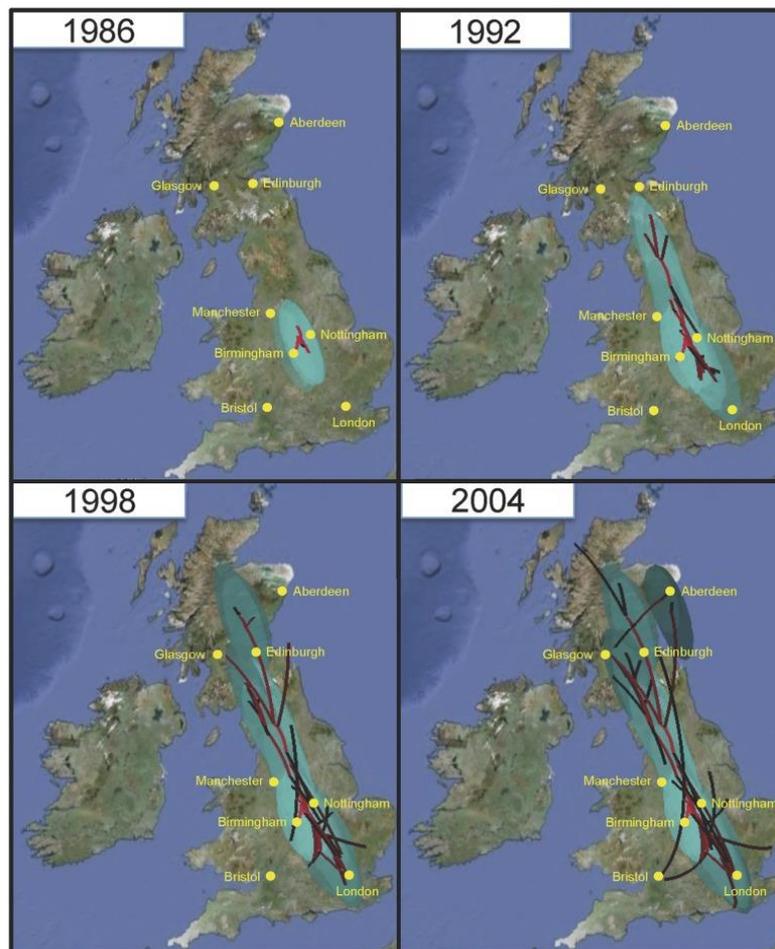
Figure 1.5: Predicted dissemination of ST22-A2 clone within the UK using Bayesian reconstruction reported from [105].

Through phylogeographic modelling, evidence suggests that the emergence of ciprofloxacin resistance among an EMRSA-15 subclone (ST22-A2) within the UK during the 1980s, originated in the West Midlands [102]. Ciprofloxacin resistance was acquired with mutations in genes encoding topoisomerase IV (gyrA) and DNA

gyrase (*grlA*) which are the primary targets of fluoroquinolones [103]. This acquisition closely resides to the introduction date and is thought to have significantly facilitated the rapid expansion and dominance of EMRSA-15 within those regions and across the UK (Figure 1.5). Throughout this period the proportion of *S. aureus* bacteraemia rates caused by MRSA increased from < 2 % to ~ 37 %, of which EMRSA-15 accounted for more than 60 % of those incidences [104].

Furthermore, phylogenetic reconstruction of 1013 MRSA genome sequences collected from the BSAC Bacteraemia Resistance Surveillance Programme between 2001 and 2010 across the UK and Ireland, discovered that 783 (77 %) isolates were members of clonal complex 22 (ST22). Over the following decades, ST22 clones have continued to demonstrate resistance to an extended range of antibiotics of various classes including tetracyclines, oxazolidinones, aminoglycosides and mupirocin on numerous occasions [102].

1.4 Bacteriophage

Bacteriophage (phage) are obligate parasites that specifically infect bacterial cells are the most abundant living entity in the biosphere outnumbering their bacterial cells ten-fold with an estimated 10^{31} viral particles present on earth [105]. Ubiquitous throughout all environments, their abundance and distribution is tied to the presence of their bacterial hosts, where they play a substantial role in regulating microbial balance and driving genetic diversity amongst bacteria.

1.4.1 Phage structure and taxonomy

Phage classification is governed by the International Committee for Taxonomy of Viruses (ICTV) based on key factors including phage morphology, nucleic acid constitution and the physiochemical properties of the virion [106]. The order *Caudovirales* (tailed phages) accounts of the majority (96 %) of bacteriophage described in the literature surpassing any other virus taxon [107], with members of this order currently characterised into three phylogenetically related families based on the morphology and function of the tail:- contractile (*Myoviridae*), long noncontractile (*Siphoviridae*) and short tailed (*Podoviridae*) [108]. These all have double stranded DNA genomes that are enclosed within a polyhedral capsid (head), a tail that acts as a passageway for the genomic material to enter the host during infection and tail fibers, that mediates the binding to the receptors on the surface of the host cell as depicted in Figure 1.6.

The ICTV classification (https://talk.ictvonline.org/) has undergone some intense scrutiny over the last two decades [109–111], the level of diversity of bacterial viruses within the order of *Caudovirales* surpasses any other organisms especially at the nucleotide level yet bacterial viruses are currently assigned in to a hierarchy of taxonomic ranks consisting of three families. With the exponential increase in prokaryotic virus genome sequences deposited on to public databases and the large majority of those currently deposited yet to be officially classified by the ITCV, this approach is no longer deemed sufficient for correctly classifying viruses as similar phage morphology does not imply genetic similarity. In 2019, the re-evaluation of current taxonomic assignment shifting towards a more genome-

based classification approach that reflects the genetic relationship of these viruses at a higher taxonomic rank was explored. The ICTV subcommittee studied the diversity of the *Spounavirinae* genus, a large group of *Myoviridae* containing notable phage members including staphylococcal phage K, G1 and Twort, as well as Bacillus SPO1, Listeria phage P100 and A511 [112]. By disconnecting the classification of *Caudovirales* based on morphotypes and family, this approach retains the virus morphotype assignment, but allows taxonomically related clades to be grouped across morphotypes. The introduction of additional ranks in to the current virus taxonomy was proposed and subsequently approved by members of the ICTV (consisting of executive committee, life members, ICTV subcommittee members (including the study group chairs) and ICTV national representatives) by an absolute majority [113]. As a result, significant changes in the ICTV Code including the addition of seven new virus families were considered. The establishment of two new family ranks, *Ackermannviridae* and *Herelleviridae* within the order of *Caudovirales*, the latter now holds the previous members of the subfamily *Spounavirinae*, which has since been removed from the family *Myoviridae* [113–115].
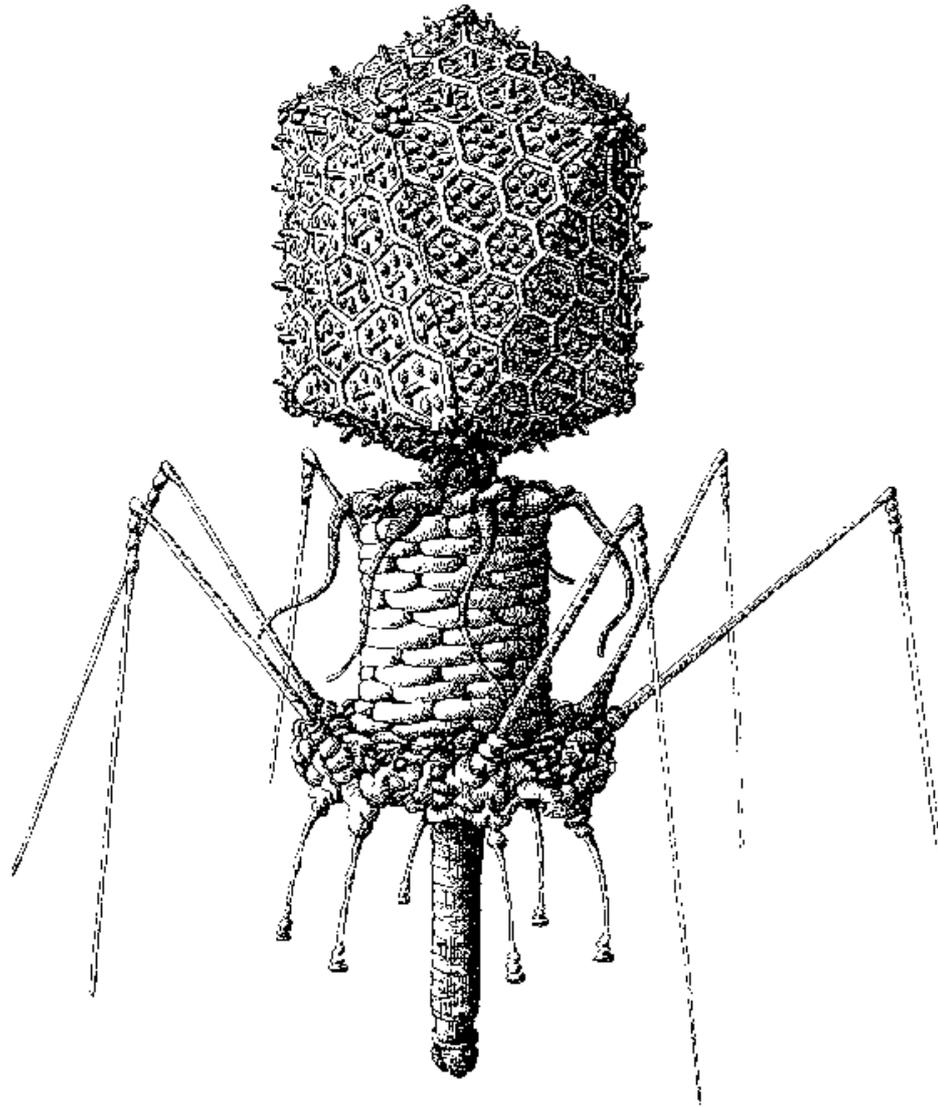
Figure 1.6: Graphical representation of a typical *Myoviridae* phage.

The icosahedral head that contains the double stranded DNA, tail fibers attached to the baseplate and the contracted tail can be identified, figure adapted from [387].

1.4.2 Phage life cycles

Phage can be differentiated depending on the life cycles that they exhibit - lytic or lysogenic, although some phage may also have pseudolysogenic component in their life cycle (Figure 1.7), which describes a phage-host relationship that neither establishes a lytic nor lysogenic response, whereby a continuous secretion of

phage occurs during a chronic infection that reduces the host cell growth rate [116]. Obligately lytic bacteriophage, adsorb to the bacterial membrane and introduce their genomic material into the host cytoplasm shutting down the bacterial hosts metabolic machinery before subverting it to resource the replication of phages own genome. This leads to the reproduction of multiple phage progeny before causing cell lysis through the use of lytic enzymes, these enzymes form holes in the bacterial cell wall, causing the cell to lyse due to the internal turgor, subsequently releasing mature virions that can go on to infect other host cells.

Temperate bacteriophage that exhibit a lysogenic lifestyle incorporate their DNA into the bacterial host chromosome to form a prophage that is duplicated and vertically transmitted to progeny cells [117]. Following integration into the host genome, prophage DNA enters a latent state (temperate) whereby expression is controlled by a repressor gene system [118,119]. The prophage switches to a lytic life cycle in response to environmental stresses that downregulates the actions of prophage repressors leading to transcription of phage genes in the first stage of the lytic pathway. The induction of prophage can be spontaneous, or induced by environmental responses caused by UV, antibiotics or stress.

a Lytic

Phage capsid
Bacterial genome
Phage genome
Capsid and tail proteins
Phage genome
Assembled phage
Lysis

b Lysogenic

Insertion
Prophage
Dormant state

c Pseudolysogenic

Non-replicating preprophage
Nutrient-rich conditions
Lytic
Lysogenic
Starved cell
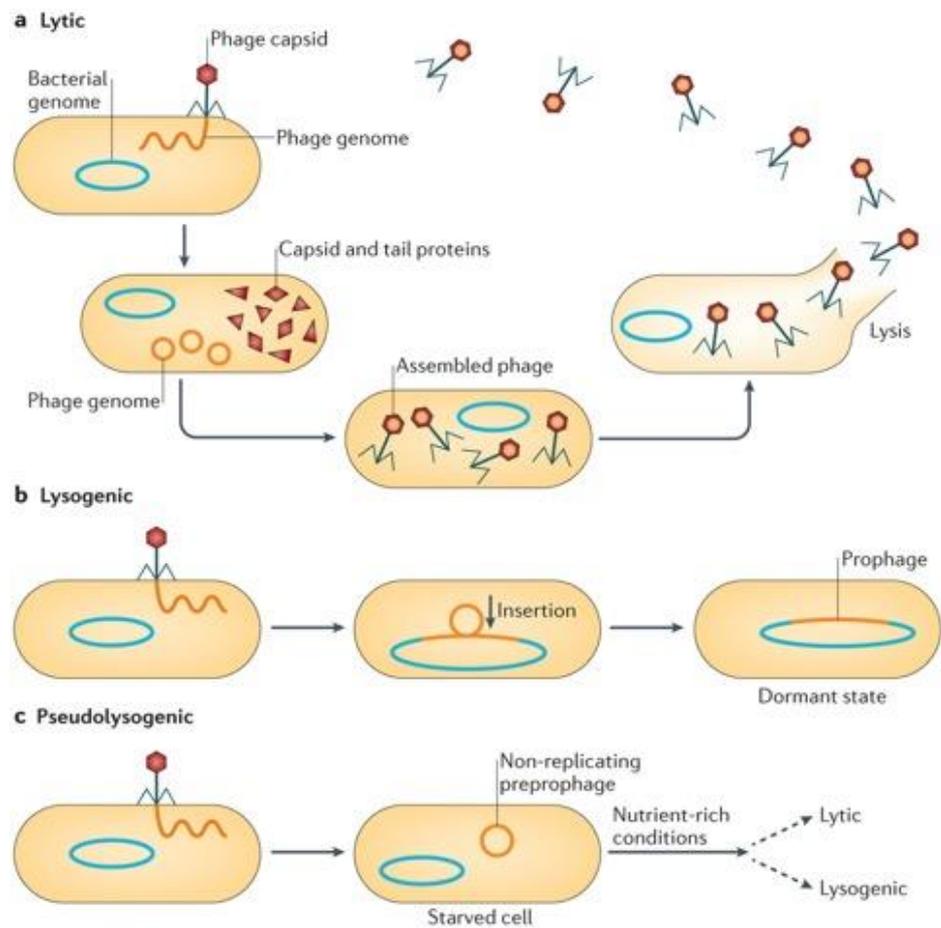
Figure 1.7: The replication stages of Lytic (A), Lysogenic (B) and Pseudolysogenic (C) phage life cycles, figure adapted from [123].

1.4.3 Phage and HGT

The lytic 'switch' can sometimes lead to improper incision of the prophage and neighbouring bacterial genes may also be packaged in virions. This acquisition of neighbouring bacterial genes known as transduction can result in the phage serving as carriers for pathogenicity islands and genes that encode for bacterial toxins and other virulence determinants [27,120]. Whilst can potentially increase the virulence of the recipient due to expression of virulence gene products during lysogeny, playing a significant role in influencing the evolution and emergence of new virulent bacteria strains [121,122]. This has been observed by the prevalence

32

of prophage carriage among *S. aureus* genomes. The success of notable CA-MRSA clones including ST8, ST30, ST59 and ST80 can be attributed to the multiple phage-encoded virulence factors including PVL and Enterotoxin A [123,124].

1.4.4 Phage evolution

Recent advances in sequencing technology in the last decade has seen an exponential increase in bacteriophage genomes published on to online databases, such as the NCBI and the European Nucleotide Archive (ENA). Other than the rare jumbo phage that harbours a significantly larger genome, the majority of phage carry comparatively smaller genomes than their bacterial hosts [125]. Despite their size and considerable abundance, phage genomes sequencing continues to lag behind bacterial genome sequencing [126]. Phage genomes display extensive mosaicism revealing genes organised into functional modules that are frequently exchanged within a population and has been described among phage belonging to *Myoviridae* and *Siphoviridae* [127]. The size of these modules and their exchange rate are highly variable [128], homologous recombination has been found to play a central role in the re-assortment of phage genomes, with both point mutations and insertion/deletion also contributing to their diversity [129,130]. Comparative genomic studies have contributed immensely to developing a greater understanding of phage diversity, their genomic architecture and evolution [127,128,131]. However, the incredibly mosaic nature of phage genomes makes detailed comparisons extremely challenging [132]. Phage mosaicism can be assessed through nucleotide sequence comparisons or at the protein level by comparing gene products, the latter approach can be applied when some phage

groups share no nucleotide similarity, and protein sequence data can reveal genes that share older ancestry [133]. They also reveal regions that are commonly conserved and those which are more variable or unique to specific phage. Importantly, despite the high levels of observed mosaicism, the analyses also identify distinct genome lineages among phage of a specific host species [131].

1.4.5 Phage lytic proteins

Phage encode for a number of lytic proteins that are used to mediate cell lysis and release viral progeny from cells. In dsDNA phage, lysis is the result of permeabilization and disruption of the bacterial cell wall achieved through the holin and endolysin system. Both holins and endolysins are produced towards the late phase of phage gene expression, during virion assembly, endolysins accumulate in the cytosol whereas holins will accumulate in the membrane [134]. Phage-encoded endolysins are multi-domain peptidoglycan hydrolases that act by digesting the peptidoglycan of the bacterial cell wall and are classed based on the catalytic specificity, such as the target substrate in bacterial cell walls [135]. These evolutionary advanced hydrolases are highly efficient enzymes that target one of the four bonds of the peptidoglycan, this activity is achieved through cleavage of the glycosidic linkages, amide or peptide cross-bridges of the peptidoglycan [136]. The folded endolysin molecules typically harbours at least two functional domains, a catalytic N-terminal domain (EAD), containing either endopeptidase, amidase, glucosaminidase, muramidase or trasnglycosylase activity. They also possess a C-terminal cell wall binding domain (CBD) which confers specificity for certain cell wall types playing a significant role in the activity range of the enzyme [137].

Endolysins do not have signal sequences therefore accumulate in the cytosol and are translocated with the help of holins in a timely controlled fashion and at a genetically specified time, the holins form lesions in the cell membrane leading to its permeablisation allowing access for the endolysins [138,139]. Cleavage of the peptidoglycan will cause the cell to lose its structural integrity leading to unsustainable osmotic pressure, ultimately bursting the bacterial cell and releasing phage progeny.

Endolysins require holins to essentially gain access to the cell wall and reach the peptidoglycan thus facilitating cell lysis terminating the infective cycle, therefore the holin is the direct determinant for the length of the phage infective cycle and effectively the burst size. The longer the latent period within a single cell, the more progeny are expected to be produced [139,140]. Endolysins are capable of causing rapid bacterial cell lysis when applied extracellularly against antibiotic-sensitive and antibiotic-resistant strains and have so far failed to instigate any form of bacterial resistance [141]. Several studies have demonstrated the antimicrobial efficacy of phage-derived lytic enzymes, including endolysins and their potential for controlling multiple Gram-positive and Gram-negative bacterial species over the past decade [142–146]. The ability to target and degrade several polymeric components of biofilm simultaneously further supports the potential of endolysins to combat biofilms. One of the most notable and well-studied endolysins is the staphylococcal phage lysin, LysK. Which has demonstrated a high lytic ability against all tested staphylococcal strains including MRSA isolates within static biofilm and animal infection models [146–148].

Virion-associated peptidoglycan hydrolases (VAPGH) are phage-encoded lytic enzymes that are involved in the initial stage of phage infection. Phage utilise VAPGHs to cleave the peptidoglycan polymers of the bacterial cell wall, facilitating the access of the phage tail tube to penetrate the bacterial cell wall, thus enabling the transfer of phage genetic material in to the host cytoplasm. Often located beneath the baseplate and locally degrade the peptidoglycan during infection [149,150], these structural components are not essential for phage replication. However their presence facilitates infection in less optimal conditions and may be used by phage to gain a selective advantage. Although VAPGHs biological characterisations are limited, comparisons of VAPGH from phages infecting S. aureus revealed that they share a high degree of similarity to endolysins at the amino acid level [151]. VAPGH have a modular structure, consisting of one or two enzymatically active domains (EADs), these catalytic domains are responsible for the cleavage of specific bonds within the peptidoglycan. Unlike endolysins, all VAPGHs described to date do not possess an identifiable cell-wall binding domain, this is likely due to that fact that VAPGHs are often delivered to the peptidoglycan by the phage particle structure [152]. This modular organisation enables the possibility of engineered domain swapping and deletion, subsequently enabling the ability to improve the lytic activity whilst also reducing the likelihood of resistance developing [151,153].

Bacteria display an array of polysaccharides on the surface of the cell such as capsular polysaccharides (CPS), exopolysaccharides (EPS) present in bacterial biofilms [154,155]. These often play a significant role in the pathogenesis of

bacteria by surrounding the cell and protecting it from desiccation, host immune responses, phagocytosis, antibiotics and phage [156]. This protective layer represents the first physical barrier phage encounter in the infection process by blocking the primary receptors phage bind to, subsequently inhibiting adsorption.

To counter this, phage encode for polysaccharide degrading enzyme known as depolymerases that recognises and enzymatically degrades the polymeric components of the EPS and capsule enabling them to reach the bacterial host [150]. Whilst playing a key role in the liberation and release of phage that are trapped within the complex structures of the biofilm EPS. The bacteriostatic activity of depolymerases produced by capsule targeting phages can be observed using soft agar overlays by the presence of a turbid zone known as a halo surrounding the phage plaque. The formation of the halo is achieved by the diffusion of soluble depolymerases through the agar and stripping away the protective capsule layer of the bacterial host. Although not necessarily killing the cell, it essentially disarms the bacteria, making it vulnerable to phage infection as well as other antimicrobials [156,157].

1.5 Phage therapy

Phage therapy exploits the natural killing ability of lytic bacteriophage and their derivatives and has been explored as a means of controlling bacterial infection. Obligately lytic phage are regarded as the most appropriate candidates in human health as they are capable of rapidly killing the target host, greatly reducing the chances of bacteria developing phage resistance [158,159]. They also lack the

required genetic factors for genome incorporation associated with temperate phage, therefore reducing the possibility of integrating their genome into the host chromosome or mobilising bacterial DNA that could potentially carry genes encoding for antibiotic resistance and virulence determinants [160,161].

Despite widespread use in the 1930's, the use of phage therapeutics were largely eclipsed by the discovery and clinical development of broad-spectrum antibiotics in Western countries. However, phage therapy has continued to be used throughout the Soviet Union with both the Eliava Institute of Bacteriophage, Microbiology and Virology in Tbilisi, Georgia and the Institute of Immunology and Experimental Therapy in Wroclaw, Poland becoming major centres for the development and application of phage therapeutics taking considerable strides in phage related research [162,163]. Which has since been integrated within their health care systems for use in the treatment of chronic and recurrent infections in patients, such as staphylococcal skin and soft-tissue infections in diabetic patients where lytic phage are matched to cultured wound organism apparently with success [164]. The efficacy of phage therapy has already demonstrated its potential when administered to patients exhibiting bacterial infections that are unresponsive to conventional antibiotics, resulting in an average success rate of around 85 % [165].

1.5.1 Advantages and limitations of phage therapy

Phage exhibit a number of advantages for therapeutic application, the mechanisms involved in phage infection are distinct from those causing antibiotic resistance

therefore antibiotic susceptible- and multiply AMR bacteria are equally as likely to be susceptible to phage killing [166,167]. Phage can encode for auxiliary enzymatic activities such as cell wall (Gram positive bacteria) and capsular polysaccharide (Gram negative bacteria) degrading properties that can disrupt the extracellular polymeric substance (EPS) matrix of biofilms, facilitating phage penetration and subsequent multiplication deep within biofilm matrices [168]. Unlike antibiotics, that are synthetic or semi-synthetic in their production, bacteriophage are naturally occurring and their ability to proliferate at the site of infection means that only small doses of phage are required. Suitable lytic phage infecting most species of pathogenic bacteria are readily isolated from environments where their hosts are plentiful such as hospital effluents, sewage or watercourses. Lytic phage vary widely in their host specificity. Examples exist of phage that are specific for receptors carried by only a small proportion of a bacterial species, however, broad-host range (polyvalent) phage are readily isolated that can infect a large number of strains within a species with some infecting strains of multiple species [169].

Despite their apparent advantages there remain several barriers to the widespread clinical use of phage in human therapy. Safety concerns exist regarding possible phage interactions with patient tissues and host microflora including potential immune responses that may limit their medium-long term efficacy [170]. Because phage generally exhibit a higher specificity for their host than antibiotics, identification of the infection-causing organism may be required to guide therapy whereas broad-spectrum antibiotics can be used empirically, in advance of knowledge of the etiology of the infection. The evolution of phage resistance in

39

bacteria occurs less frequently than for some antibiotics, however this can be addressed through the combined use of different phage targeting the same bacterial strain that require a number of independent mutations to confer resistance – analogous to the use of antibiotic combination [163]. The development of phage resistance and the limited host range of individual phage can be ameliorated by the rational formulation of combinations of multiple polyvalent phage. These 'cocktails' can enable the targeting of multiple strains and even species while greatly reducing the probability of bacteria developing phage resistance [159,171]. In order to satisfy regulatory requirements in human medicine it is essential to generate high purity phage lysates that satisfy existing regulations with regards to contaminants that may act as adventitious agents and pro-inflammatory molecules such as exotoxins [172,173]. Complete DNA sequencing of phage components is desirable in candidate therapeutics to ensure they do not contain lysogenic sequences such as integrase genes, or host DNA such as bacterial virulence genes [27].

1.6 Bacteriophage resistance

Bacteria have developed a number of anti-phage mechanisms as a response to defend and prevent the constant threat of phage infection (Figure 1.8). Despite this however, phage also have a number of adaptive strategies that allows them to circumvent this issue, and it is this evolutionary arms race between hosts and phages, that is one of the greatest drivers in genetic variation and diversity.

1.6.1 Restriction Modification

A major mechanism of resistance to bacteriophage infection are bacterial restriction-modification (R-M) systems. Characteristically, these consist of a number of enzymes including restriction endonucleases (REase) and DNA methyltransferases (MTase) that rapidly cleave and methylate foreign DNA at specific sites upon entry to the cell, deactivating it prior to phage DNA transcription [174]. R-M systems are typically classified in to four major types ($I - IV$) that are based on their subunit composition and number, enzymatic mechanism, cleavage position, sequence recognition and substrate specificity [175]. R-M systems are an extremely diverse group of enzymes and are widespread throughout prokaryotes, the variation and prevalence of this system indicates the success and importance as a defensive mechanism within the bacterial world. Due to this diversity, phage have evolved a number of anti-restriction strategies to counter and evade R-M systems. Such mechanisms include the incorporation of modified bases, altering the restriction site and reduction of endonuclease recognition sites through the accumulation of point mutations within their genomes [174].
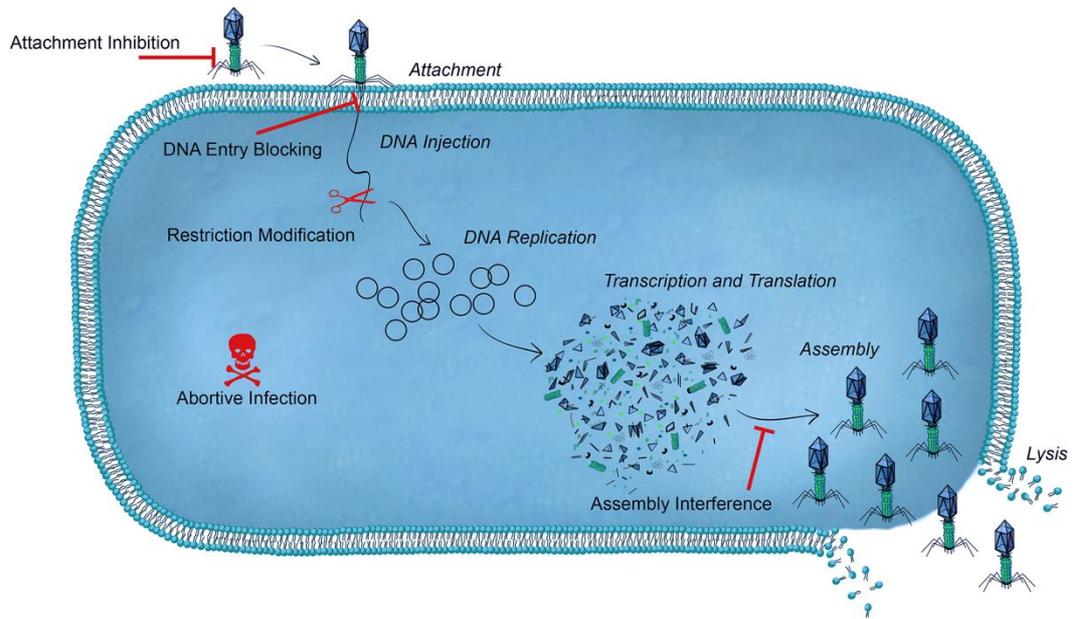
Figure 1.8: The various bacteriophage resistance mechanisms employed by bacteria to target different stages of the phages lytic life cycle. Figure adapted from [182].

1.6.2 CRISPR-Cas system

Clustered regularly interspaced short palindromic repeats (CRISPR) as well as CRISPR-associated (Cas) proteins form part of the adaptive and heritable immune system within prokaryotes against mobile genetic elements such as phages and plasmids [176]. The system consists of several short palindromic repeats that are identical in length and sequence, these conserved repeats are interspaced by variable DNA segments called spacers derived from previous exposure to phage and other MGEs. The sequence of each spacer is 100 % identical to the spacer sequence found in the genome of the infecting phage known as the protospacer and enables the host to recognise and destroy specific phages acting as a resistance mechanism against foreign genetic elements [177]. Spacers are also flanked by a varying number of *cas* genes, that when transcribed, form an array of complex cas proteins with functional domains and are characteristic of nucleases,

helicases, integrases and polymerases [178]. The CRISPR arrays (including the foreign DNA spacers) are then transcribed in to small non-coding interfering CRISPR RNAs (crRNA), which direct the Cas proteins to specifically target and cleave the foreign genetic material [174,178].

The CRISPR-Cas system is widespread among bacteria and archaea, with most possessing only one CRISPR locus. However there are several bacterial species that contain multiple CRISPR loci [179]. To overcome this, phage resistance to the CRISPR-Cas system can be acquired through the presence of anti-CRISPR (*acr*) proteins that inhibit a range of Cas9 proteins. However, even a simple point mutation or deletion in the protospacer, or the protospacer adjacent motif (PAM). The PAM is a short (2 − 7 bp), conserved sequence of DNA located adjacent to the protospacer within the infecting phage genome, that is targeted by the Cas9 nuclease can be enough to grant immunity and evade this defensive strategy [180].

1.6.3 Adsorption blocking

Bacteria present an array of host-specific components on the cell surface that play a significant role in bacterial metabolism but also the virulence and pathogenicity of the host. However these also serve as the key receptors that phage target and irreversibly bind to during the initial stage of infection in order to inject the genomic DNA through the host membrane [181]. Adaption to the structure of the surface components, production of extracellular matrix such as capsule or glycocalyx layer, or by competitive inhibitors that bind to phage receptors are some innate defensive strategies bacteria employ to inhibit phage infection [174,182].

Receptor availability can be greatly decreased through phase variation, this heritable, but reversible process regulates gene expression enabling population heterogeneity in attempt to survive phage attack. Alternatively, such receptors can be physically masked by the production of exopolysaccharides to surround the bacterial cell in a slime or capsule preventing phages attachment [183]. To counter this, some phage encode a number of polysaccharide-degrading enzymes that are capable of breaking down the exopolysaccharide layer thus revealing the host-specific receptors to initiate phage infection. These enzymes are covered in section 1.4.5.

1.6.4 Abortive infection

The abortive infection (Abi) system is an altruistic action that promotes the death of the phage-infected bacterial cell as a way of sacrificing itself, in attempt to protect the surrounding bacterial population from subsequent infection. This system can act at any stage of phage development by targeting the essential cellular processes within the host such as replication, transcription and translation in attempt to reduce or eliminate the production of phage progeny [183]. It has been found that Abi systems are often encoded on MGEs, and can be mediated through toxin-antitoxin systems, that are widespread within prokaryotes [184]. In order to circumvent this, some phage are capable of encoding for their own antitoxin molecules that are used to suppress the bacterial toxins required for Abi to occur [183].

1.6.5 Bacteriophage assembly interference

Gram positive bacteria possess phage-inducible chromosomal islands (PICIs) that parasitise and interfere with the reproduction of infecting phage. One of the most well studied PICI is the *Staphylococcus aureus* pathogenicity islands (SaPIs) that encode for a number of virulence factors and contribute substantially to host adaptation and pathogenicity as described in section 1.2.5.1.

Normally, SaPIs resides stably within the bacterial chromosome but can be mobilised by the presence of 'helper' phage through infection or lysogeny induction, subsequently leading to the excision, replication and packing of the SaPI. All known SaPIs described so far target essential phage functions and interfere with helper phage particle assembly and DNA packaging using several strategies [182]. They can remodel the capsid proteins of the infecting helper phage so that its smaller and more tailored to fit the SaPI genome. SaPIs also encode for phage packaging interference (Ppi) proteins that diverts phage packaging towards the SaPI by blocking the terminase small subunit of the helper phage, but not the corresponding SaPIs terminase small subunit [185]. A third mechanism involves interfering with late phage gene transcription, which is essential for phage packaging and cell lysis, however must be modulated precisely so that it does not also interfere with SaPI particle formation [186].

1.7 Phage effects on biofilms

The therapeutic potential of bacteriophage is becoming increasingly explored over recent years, several studies have claimed promising success in the effective disruption and dispersal of biofilms produced by clinically relevant pathogens [187–191]. The extracellular matrix not only provides structural integrity but also acts as a protective barrier around the bacteria that increases their tolerance to antimicrobial agents and innate host defence mechanisms [192,193]. In a clinical scenario, biofilms present a major obstacle in the treatment of bacterial infections using conventional chemotherapeutic methods, requiring alternative strategies for eradicating or reducing complex microbial communities [194].

Within the anoxic conditions of a natural biofilm environment, the high local concentration of cells are exposed to a myriad of threats such as depleted nutrients and antibiotics produced from competitors. This results in the programmed formation of persister cells that are metabolically dormant [195]. Due to this state of dormancy, persisters display a higher tolerance to antimicrobials that are mainly effective against actively growing cells by targeting metabolic pathways and are difficult to eliminate, resulting in waves of biofilm regrowth following disinfection procedures [39,196]. The down regulated metabolic machinery of persister cells required by phage for intracellular replication inhibits the completion of the phage infection cycle. However It has been previously reported that phage are still capable of adsorption to these persister cells, although no replication will occur, they too will enter a quiescent state within the cell until the host becomes metabolically active allowing the phage infection cycle to begin [197].
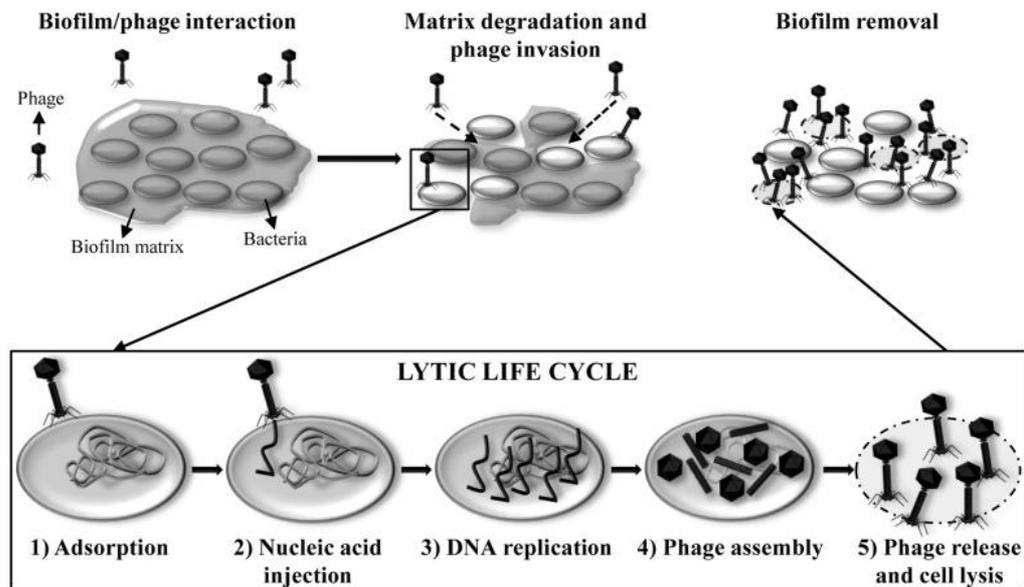
Figure 1.9: Lytic phage infection within biofilms. Figure adapted from [64].

Phage are equipped with a number of other properties that make biofilms susceptible to their action [135,190,198]. Although larger than chemical antibiotics, phage are much smaller than their bacterial hosts and have the ability to penetrate deep within the protective layers of biofilms [168,194]. Water-filled channels run through the structure of mature biofilms in which nutrients can circulate, such channels are large enough for phage to pass through and invade the interior and basal layers of the biofilm [169]. Their ability to self-replicate within a host resulting in localised amplification of phage progeny that can go on to infect many more hosts, allowing them to reach bacterial cells that were once protected and out of reach (Figure 1.9). To circumvent such hurdles, phage-derived enzymes with antibacterial properties could be introduced as antimicrobials. Several studies have examined their ability to disrupt biofilms by targeting the key components of the glycocalyx that makes up the majority of the biofilm structure causing large

masses to slough off, resulting in near complete dispersal of the entire biofilm [199].

1.8 Staphylococcal phage

Staphylococcal phage are widespread and have been extensively studied for many decades. Phage typing is one approach that utilises the high specificity of phages to successfully subtype *S. aureus* isolates, as a way of characterising the epidemiology of clones, especially those responsible for causing outbreaks of infectious disease [200]. Another approach is to utilise the lytic ability of phage as a means of treating multiply antibiotic resistant bacterial infections for over almost a century [172]. Staphylococcal phage in therapeutic use e.g. Eliava institute in Tbilisi, Georgia, are obligately lytic, polyvalent phage capable of infecting multiple strains of a *S. aureus* including MRSA and coagulase-negative *S. aureus* – CONS [201]. The most effective tend to be Twort-like phage myoviruses and lytic podoviruses that exhibit short eclipse times and high burst sizes [187,202–205]. To aid in the attachment, phage utilise a number of adhesins that recognise cell surface receptor proteins of the host cell. The specificity of these adhesins defines the host range of the phage. The adsorption of phage particle to the host cell is deemed to be the most important determinant for successful infection. Phage can infect host bacteria with great specificity, where they are capable of infecting bacteria from a single species, and in some cases a single strain. However, some phage do exhibit a wide host-range (polyvalent) with the ability to infect the majority of strains of a single species and in some cases even multiple species [206]. Host recognition is achieved through binding to molecules presented on the

host cell wall through the use of phage receptor binding proteins [207]. Such molecules include wall teichoic acid (WTA) polymer, which is abundantly found on the surfaces of *S. aureus* and are major constituents of their cell envelope [208]. The composition of these glycopolymers of backbone of WTA varies and can be lineage-specific, however the majority of *S. aureus* strains harbour a polyribitol phosphate (Rbo-P)-type WTA. Several studies have indicated that the N-Acetylglucosamine (GlcNAc) residue of WTA is a key target for phage infecting *S. aureus* playing a pivitol role in host recognition and defining the host range of the phage [209,210].

The renewed interest in phage therapy has led to an increase in a variation of novel *S. aureus* phage isolated, with several recent studies describing the lytic activity of staphylococcal phage in the treatment of *S. aureus* and *S. epidermis* using experimental biofilm models [158,187,188,190,211]. The ability to prevent the formation and disperse pre-formed staphylococcal biofilms using phage cocktails has also been successfully demonstrated, whilst greatly reducing the rate of phage resistance [187,189,212]. The combinational use of phage could facilitate the adsorption and infection of polymicrobial biofilms, the disruptive ability of one phage against the glycocalyx structure of the biofilm through the use of lytic enzymes and depolymerases, can allow other phage within the cocktail to reach its bacterial host [213].

1.8.1 Phage antibiotic synergy

Studies investigating the combined use of phage and their derivatives with antibiotics and their activity on biofilms are still in their infancy, however previous studies have already identified possibly synergies resulting in enhanced biofilm disruption in vitro when administered in combination, or sequentially [214–217]. A limited number of studies have approached the phage-antibiotic synergy (PAS), whereby sub-lethal concentrations of antibiotics may also stimulate the host production of virulent phages. Authors suggest that pre-treatment of *S. aureus* biofilms with phage yielded the greatest results, by disrupting the bacterial biofilm structures that were once protecting cells, it allows for deeper penetration of antibiotics leading to significant reductions in bacterial viability (Figure 1.10)[218,219].
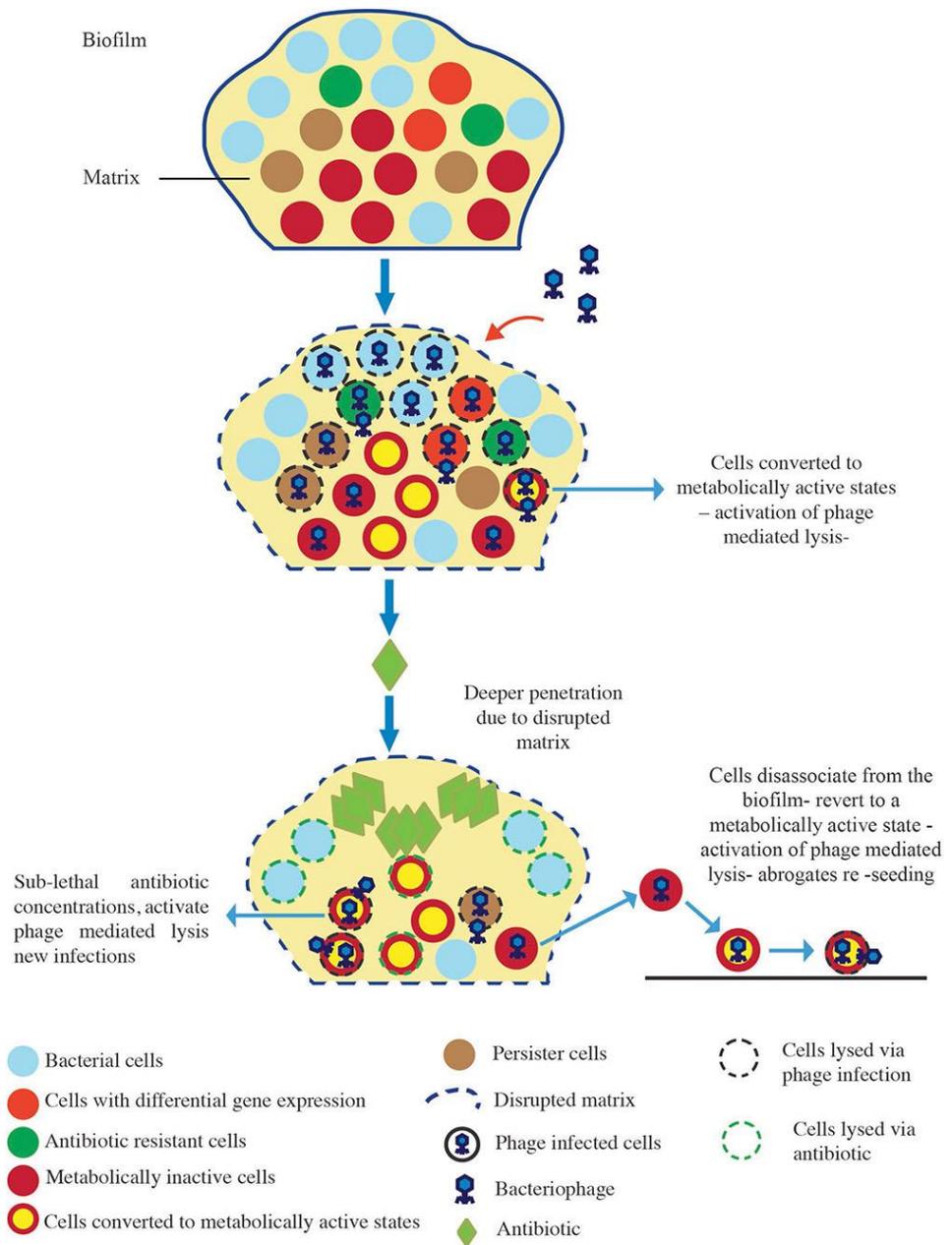
Figure 1.10: Sequential treatment of mature biofilms.

Pre-treatment of biofilm with phage preparations in attempt to disrupt and disperse the extracellular polymeric substance of biofilms, thus allowing sub-inhibitory concentrations of antibiotics to activate phage mediated lysis of previously dormant cells. Figure adapted from [218]

1.8.2 Staphylococcal phage classification

Until late 2018, *Myoviridae* infecting *S. aureus* were classified in to the *Spounavirinae* subfamily which consists of six genera including *Kayvirus*, *P100virus*, *Silviavirus*, *Spo1virus*, *Tsarbombavirus* and *Twortvirus* and three unassigned species (*Enterococcus* virus phiEC24C, *Lactobacillus* virus Lb338-1 and *Lactobacillus* virus LP65) [112,220]. The *Spounavirinae* subfamily, initially proposed in 2009 are a group of phage that are related to the *Bacillus* phage SPO1 including notable phage such as staphylococcal phage Twort, staphylococcal phage K, staphylococcal phage G1, Listeria phage P100, and *Listeria* phage A511 [112]. The characteristics of these members exhibit a strictly virulent lifestyle, myovirion morphology, possessing terminally redundant, non-permuted dsDNA genomes that are typically organised in to functional modules. The phage share considerable amino acid homology, representing a distinct cluster within the dsDNA viruses with largely overlapping sets of genes [131,220]. The reorganisation of the current ICTV classification system led to the recent establishment of a two new families named *Ackermannviridae* and *Herelleviridae* [115,131]. As a result, *Herelleviridae* now contains the four new subfamilies within it including the previous *Spounavirinae* and the new *Bastillevirinae, Brockvirinae, Jasinkavirinae* and *Twortvirinae* subfamilies (Figure 1.11). Further to this, the genera (*Kayvirus*, *Sepunavirus*, *Silviavirus*, *Twortvirus* and unclassified *Twortvirinae*) containing staphylococcal phage now reside in the subfamily *Twortvirinae* [131]. As of June 2019, there are currently 55 staphylococcal phage genomes currently available on the NCBI database as seen in the taxonomy browser.
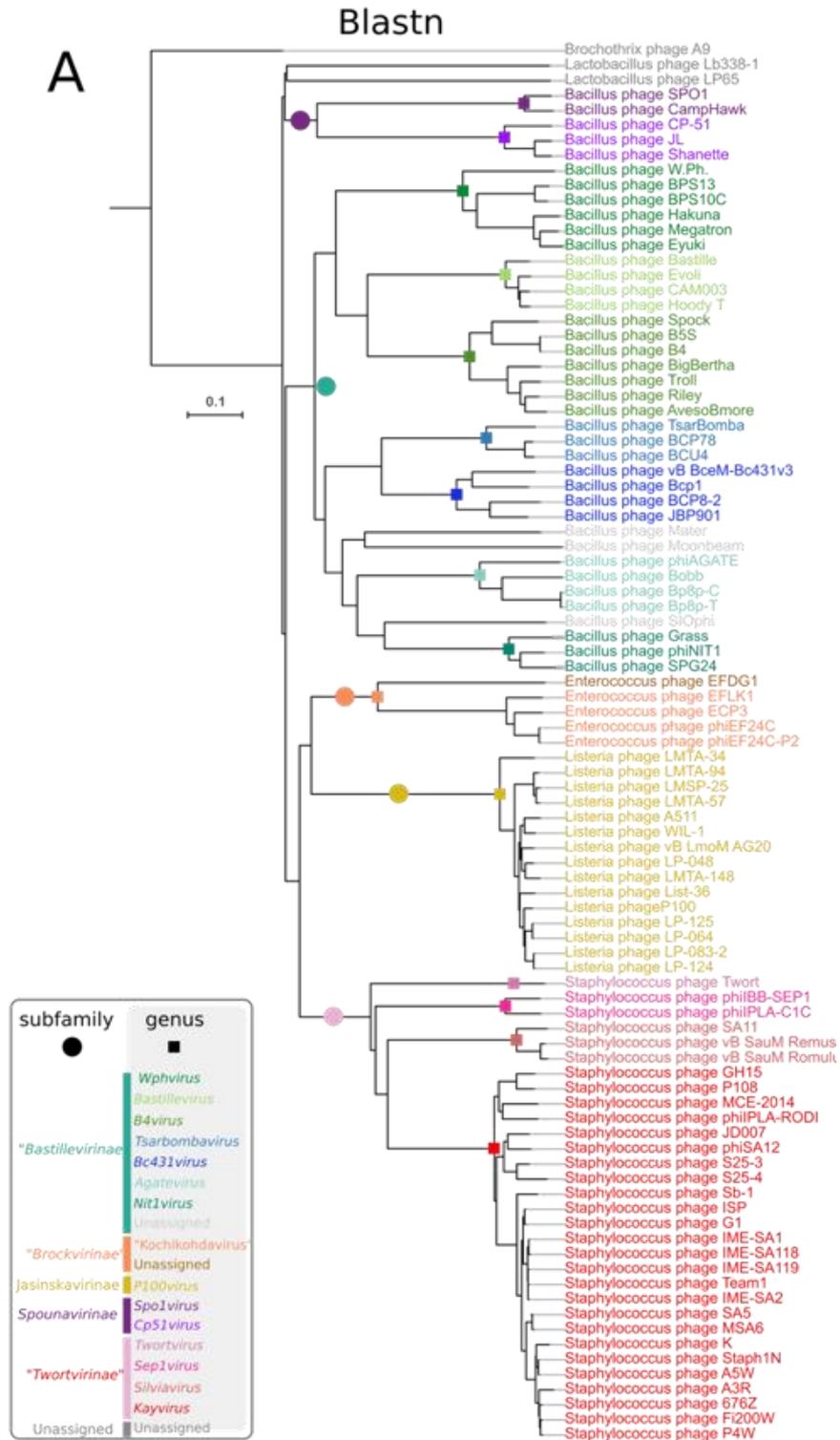
Figure 1.11: BLAST-based clustering of the newly established *Herelleviridae* subfamily.

Clusters are based on original names of each genera representation by colour heatmap generated with Gegenees. Phylogeny adapted from [131].

1.9 Aims and objectives

The main aim of this study was to develop a greater understanding of lytic bacteriophage effective against *S. aureus* and their potential to control biofilms produced by two dominant MRSA clones within the UK, ST22 and ST36. This study is divided in to three main objectives that are:

- To establish a collection of lytic staphylococcal phage isolated from environmental samples and the determination of their host range against a diverse collection of clinical *S. aureus* (including MRSA) isolates.

- To evaluate the disruptive anti-biofilm properties of candidate phage against a number of *S. aureus* biofilms produced by ST22 and ST36 isolates.

- To examine the associations between the genomes of the phage collection and other lytic phage infecting *S. aureus* using comparative genomics.

Chapter 2

Materials and methods

2.1 Bacterial strains

A total of 185 genetically diverse (based on multilocus sequence typing - MLST) *Staphylococcus aureus* isolates were used in this study (Appendix A). The isolates, from human carriage and disease, include some from the major methicillin-resistant *S. aureus* (MRSA) clones, including 43 multilocus sequence type (ST)22 and 27 ST36 from the UK and USA [77,80,102]. Additionally, a *S. carnosus* strain TM300 kindly provided by Dr. Guoqing Xia (University of Manchester, UK) was used for phage isolation and propagation. *S. carnosus* is a non-virulent staphylococcal species used in meat production and as such was considered a benign propagating host for phage production. This modified TM300 contains a plasmid encoding the minimum number of genes required for biosynthesis of *S. aureus* polyribitol-phosphate (RboP) wall teichoic acid using antibiotic-resistance gene markers. The use of an antibiotic-resistance gene marker requires chloramphenicol supplemented media at a concentration of 10 μg/mL. and is necessary for expressing the polyribitol-phosphate (RboP) repeating units of the wall teichoic acid, thus aiding phage adsorption and should increase the chances of isolating novel phage [210].

2.1.1 Maintenance and standardisation of bacterial cultures

All strains were sourced from freezer stocks maintained in 25 % (v/v) glycerol in TSB at -80 ˚C. *S. aureus* isolates were plated on Tryptone Soy Agar (TSA) and incubated overnight at 37 ˚C in air. For liquid cultures individual colonies were picked and inoculated into 20 mL Tryptone soy broth (TSB) and incubated overnight 37 ˚C with shaking at 150 rpm. TSB-soft agar containing 0.6% Agar

Bacteriological (Agar No. 1) was used for bacteriophage propagation and plaque assays. Molten soft agar was stored in a water bath at 50 ˚C before use.

2.1.2 Normalisation of cells

Bacterial cultures were standardised to the same concentration at the required optical density (nm). For this, liquid cultures were harvested by centrifugation for 20 minutes at 4200 rpm, spent medium was decanted and bacterial pellet was resuspended in PBS using the pipette to break up the pellet. Cells were centrifuged again and washing of cells was further repeated once more, cells were this time resuspended in 5 mL of TSB to create a cell concentrate. Cells were diluted accordingly to normalise all cells to required optical density. Once the desired concentration of bacterial cultures was obtained, 10-fold serial dilutions were performed (10^2 - 10^9) in PBS and 100 µL was spread onto TSA plates in triplicate and incubated overnight at 37 °C. The plate dilution displaying 30 – 300 individual colonies were used to calculate the CFU/mL at each absorbance using the following equation:

$$CFU / ml = \frac{Number\ of\ colony\ forming\ units}{Dilution\ factor\ x\ Volume}$$

2.2 Sample collection

2.2.1 Collection of environmental samples

Animal faecal samples were collected from various locations of fields holding grazing livestock, in addition to several animal pens at Chelford cattle market. Holding pens were immediately cleaned once livestock were relocated, especially cattle pens which were essentially large open rooms almost continuously sprayed down turning samples in to a mixed slurry. Solid animal faecal samples of similar size were collected in sterile falcon tubes.

2.2.2 Collection of Wastewater samples

Wastewater treatment samples were initially collected from three different stages of the treatment process - inlet, primary sedimentation and aerated sludge process tanks. Using a PTFE dipper with a 500 mL beaker, wastewater samples were collected from the surface, the dipstick was also submerged 30 cm. Around 500 mL were collected and poured into sterile sealable tubs, individually sealed again in food bags and then immediately transported to be processed.

2.2.3 Enrichment of animal samples for phage isolation

Solid animal faecal samples were homogenised using 15 mL of PBS, with aggressive agitation using a vortex mixer every 10 minutes for one hour. About 20 mL of slurry was also collected from slurry drainage channels and combined with PBS so all animal samples were roughly the same viscosity. The homogenate was centrifuged at 4200 rpm for 30 minutes and the supernatant was filter sterilised using a 0.22

μm Acrodisc syringe filter (Pall Corp., Mississauga, Ontario, Canada), to give a crude viral lysate. These lysates, from environmental and sewage samples, were enriched with staphylococcal liquid cultures to amplify phage numbers. Roughly, 5 mL of filtered sample was enriched with 20 mL of TSB and a mixed culture consisting of single or mixed cultures of actively growing bacterial cells. Enrichments were incubated overnight at 37 °C whilst shaking at 150 rpm, these were then centrifuged at 4200 rpm for 30 min to remove bacteria. The supernatant was passed through a 0.22 μm filter syringe to create an enriched phage lysate.

2.2.4 Enrichment of wastewater samples

Sewage effluent samples were collected from various process tanks at Davyhulme and Eccles wastewater treatment works, England. Organic matter was removed from samples by centrifugation (4,200 rpm, 30 mins). 10 mL aliquots of supernatant were filtered (0.22 μm), before being combined with 10 mL double-strength TSB and 100 μL of exponentially growing *S. aureus* cultures, followed by incubation at 37 °C, shaking at 150 rpm for 24 h. Bacterial debris were removed by centrifugation (4200 rpm, 30 min) filter sterilised (0.22 μm) to give an enriched phage lysate.

2.3 Phage methods

2.3.1 Soft agar overlay plaque assay

Plaque assays are performed in order to identify the presence of phage and determine their titre using the soft agar overlay method [187]. This is fundamental to a number of methods that involve phage isolation and propagation. Soft agar overlays create a confluent lawn allowing to locate areas where phage a present based on zones of lysis known as plaques [221]. 100 µl of overnight liquid bacterial cultures were aliquoted out in to 4 mL single use sterile culture tubes, 100 µl of phage lysate was added to each tube and allowed to sit at room temperature for 5 min. Working with one tube at a time, 3 mL of molten soft agar was added to the tube and immediately poured on to the surface of a TSA plate (Figure 2.1). Plates were left at room temperature for 15 min to allow the agar overlays to solidify before inverting and incubating overnight at 37 ˚C. *S. carnosus* strain TM300 was used for phage propagation whenever possible however for some phage *S. aureus* isolates D329 and H402 were used.

2.3.2 Single plaque purification

Single isolated plaques were picked by touching a sterile toothpick or pipette tip into the plaque surface and this was then agitated in 1 mL of SM buffer. Harvested phage were vortexed for 30 seconds and serially diluted before being examined for purity using the soft agar method. Single isolated phage plaques and three rounds of plaque purification were used for each phage isolate to ensure purity of lysates.
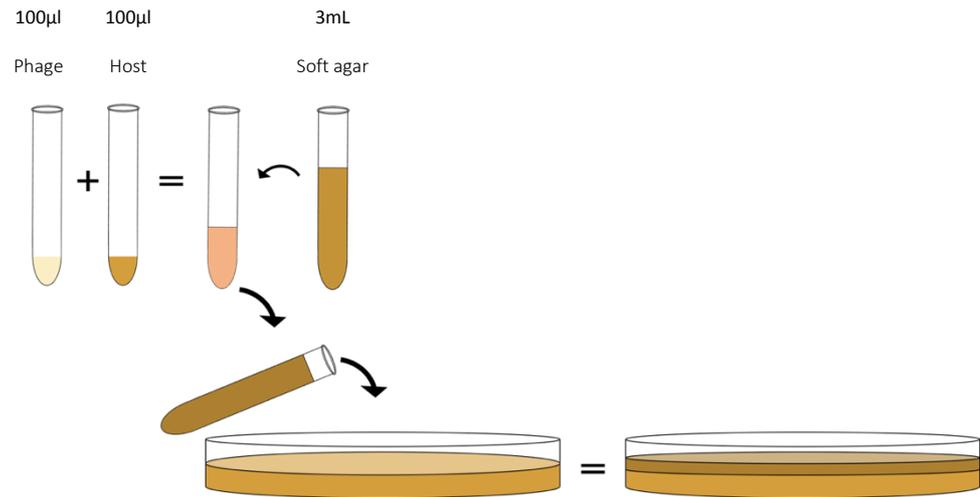
Figure 2.1: Schematic representation of soft agar overlay.

2.3.3 Phage propagation

Concentrated phage stock lysates were created using a 'flood plate' method from purified phage lysates (above). To do this dilutions of the phage suspension were plated out using the soft agar overlay method to determine the ideal dilution that produced near confluent lysis of the bacterial lawn (Figure 2.2).

The plate typically had a 'spiderweb' effect with very little delineation between each plaque. This near confluent has been found to be ideal for producing reliably high phage titers [222]. Ten plate replicates of each phage using the required dilution were created, and 5 mL of SM buffer was added to the surface of each plate before incubation at 4 °C overnight. Liquid was removed from the plate with a Pasteur pipette and transferred to a 50 mL falcon tube. These were centrifuged and the supernatant sterilised by filtration (0.22 µm pore size) to create high titre phage lysates, one or two drops of chloroform was added to each lysate and were then stored at 4 °C for short-term use.
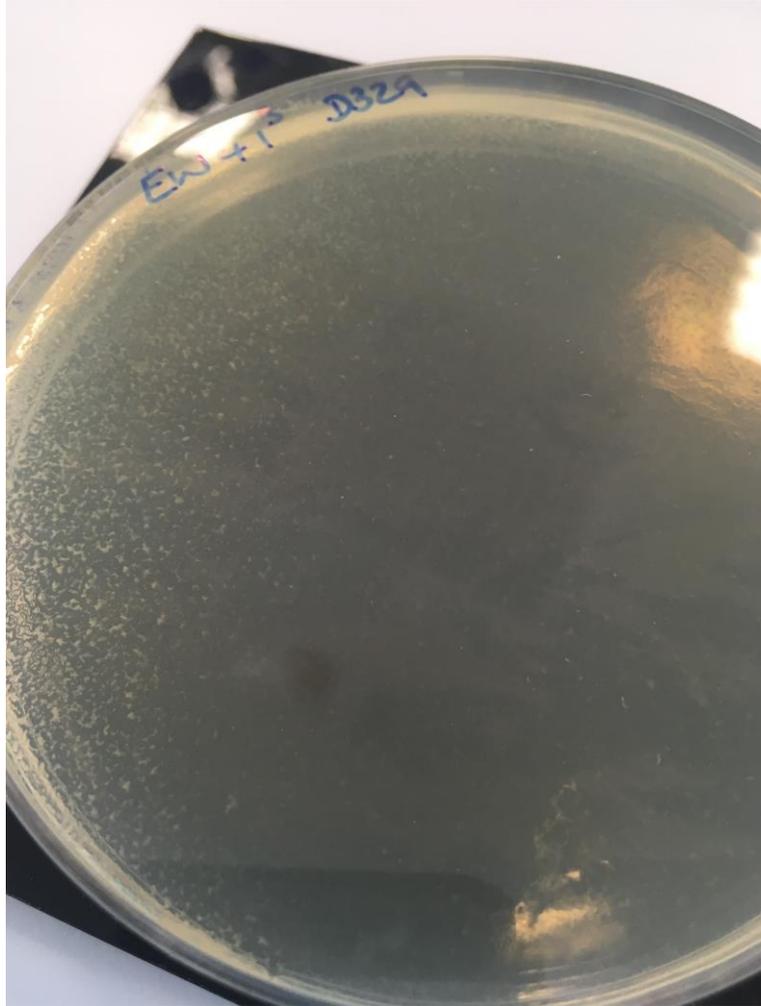
Figure 2.2: Confluent lysis of S. aureus D329 lawn. Phage dilution causing near confluent lysis were used to produce high titre phage lysates.

2.4 Phage quantification

The titre for each phage lysate was determined by the soft agar overlay. Using SM buffer as diluent, 10-fold serial dilutions of lysates (10^4 - 10^9) were prepared, 100 µl of each dilution was added to 100 µl of actively growing host strain and the soft agar overlay assay was performed, this process is performed in triplicate to obtain an average. The plates were examined for plaque presence following overnight incubation at 37 °C. if too high of a dilution is plated, confluent lysis is achieved on each plate and counting plaques is not possible. For each phage lysate, the number of plaque forming units per millilitre (PFU/mL^{-1}) were calculated the plate dilution displaying $30 - 300$ plaques whilst taking into account the dilution. The PFU/mL^{-1} is used for determining the titre based on the following equation:

$$\text{PFU}/ml = \frac{Number\ of\ plaques}{Dilution\ factor\ x\ Volume}$$

2.5 Phage host range

Prior to starting, 4 x 4 grids were drawn on the reverse side of each TSA plate to help with interpreting results and act as a guide when performing the spot assay (Figure 2.3). Overnight cultures of 185 genetically diverse (based on MLST) *S. aureus* strains (including MRSA) were prepared in 10 mL TSB (Appendix A). Soft agar overlays containing 100 µl of *S. aureus* and 3 mL of molten soft agar were mixed in sterile culture tubes and immediately poured onto the surface of TSA plates. These were allowed to set by leaving for 15 min at room temperature. 10

μl of each phage lysate, normalised to a titre of 10^6 PFU/ml, were spotted onto the surface of each plate, and allowed to dry before overnight incubation at 37 °C. Three plates were used for each phage lysates to give three replicates.
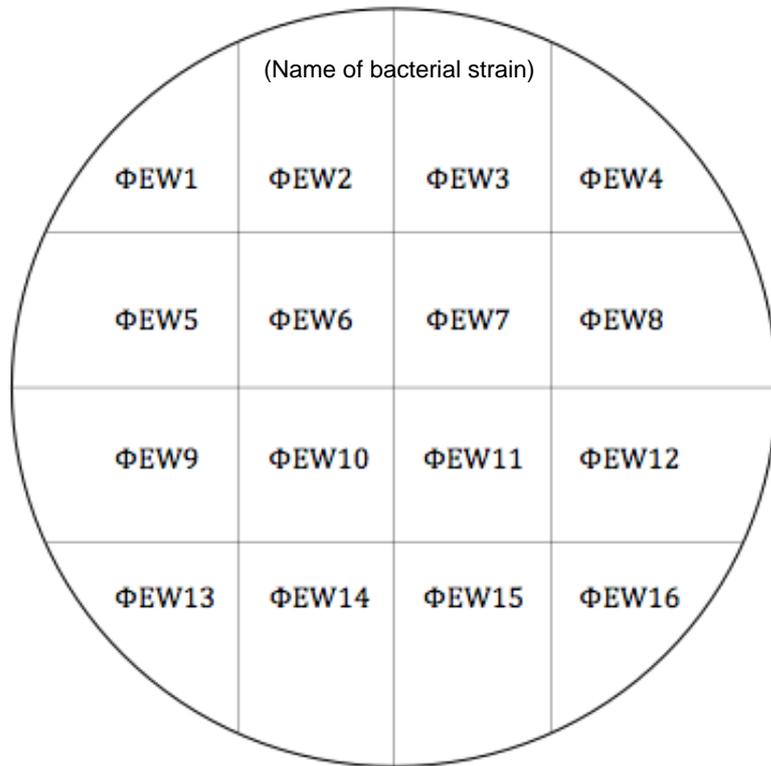
Figure 2.3: Representative plate layout used for host range assay.

The lytic ability of the phage was assessed by analysing the clearing of bacterial lawns following the spot test method. Sensitivities are also assigned a colour to simplify the analysis of the recorded results:
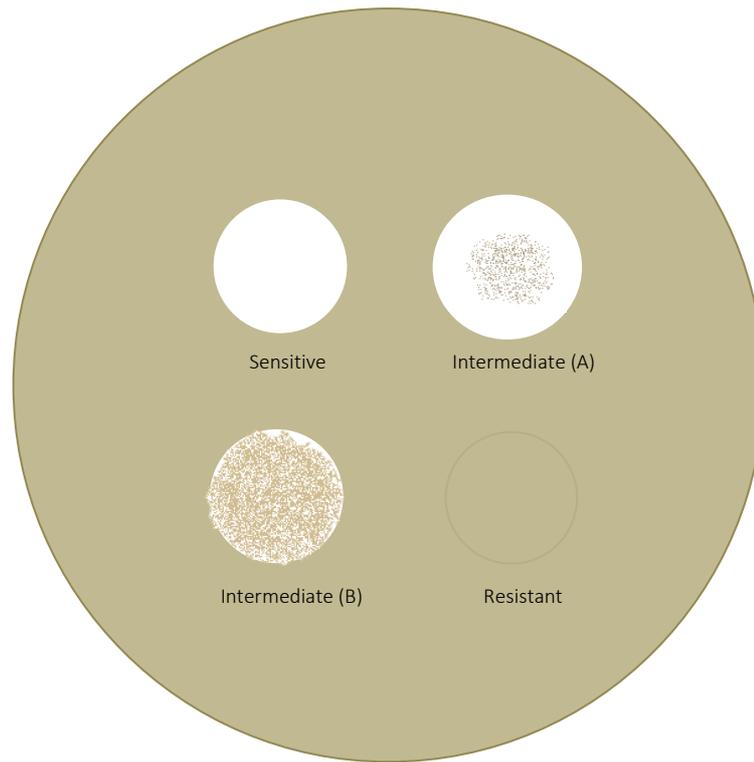
Figure 2.4: Schematic example of the degree of clearing of bacterial lawns following phage spot assay used to determine phage lytic ability against *S. aureus* isolates.

- ▬ - Sensitive (green) – completely clear spot with 100 % lysis.

- ▬ - Intermediate (yellow) – varies from clearing throughout with resistant mutants (1), clearing throughout but hazy background (2), substantial turbidity where phage was spotted (3), individual plaques (4).

- ▬ - Resistant (red) - no clearing or disruption to bacterial lawn.

2.6 In vitro growth experiments

2.6.1 Growth kinetics of bacterial hosts

The growth rate of each bacterial isolate in liquid culture was studied in 96-well flat-bottomed microtitre plates, bacterial growth over time was assessed by measuring the turbidity of each well by absorbance (OD_{600}). Initially, overnight cultures of actively growing cells were normalised to an optical density (OD_{600}) of 0.3 which corresponds to roughly 1-5 $x10^6$ CFU/mL, similarly to as mentioned in section 2.1.2. In a 96-well microtitre plate, 200 µl of sterile PBS was placed in each of the surrounding wells at the edge of the plate to help reduce sample evaporation. A 1:100 dilution of an overnight bacterial culture was performed by adding 2 µl of normalised bacterial suspension to 198 µl of fresh pre-warmed TSB, 200 µl of TSB was added to a set of wells as a negative control. All wells were replicated in triplicate. Prior to incubation, the underside of each microplate lid was sprayed with 3 % (v/v) Triton X-100 in absolute ethanol to reduce condensation on the lid that could affect measurements. Bacterial growth was measured by absorbance (600 nm) over 19 hours at 37 ˚C using a microplate reader (FLUOstar Omega, BMG LABTECH). The plate reader provided absorbance data points every 180 seconds following a 10 second agitation at 200 rpm. Average absorbance readings were standardised against blank controls, to allow determination of standard deviation / standard error for graphing purposes. Data points after every 30 minutes were used.

2.6.2 Killing assays

Time-kill assays were performed in order to determine the sensitivity of planktonic bacterial cells to phage infection and also to investigate the frequency of phage resistant bacterial mutants. Exponentially growing bacterial cultures were normalised and diluted using fresh pre-warmed TSB to an optical density of 0.300 at 600 nm. Assays were performed in a 96-well flat-bottomed polystyrene microtitre plate and these were replicated in four wells. 200 μl of sterile PBS was placed in each of the surrounding wells at the edge of the plate and 150 μL of normalised bacterial culture was added in to each test well. Lysates of selected phage were normalised to the same titre in TSB media, 50 μl was subsequently added to each well to achieve a multiplicity of infection (MOI) of 0.1 for each bacterial host. Two hundred microlitres of TSB was added to four wells as a blank control, 50 μl of sterile TSB was added to 150 μl of bacterial culture as a negative control. Microtitre plates were incubated at 37 ˚C in a plate reader and optical density was measured at 600 nm every 180s for 19 hours, experiments were performed on three separate occasions.

2.6.3 Formation of mature biofilms

The method used for studying biofilm formation on cell-culture 96-well flat-bottomed microtitre plates was based on the above method with some modifications [187]. 10 mL overnight liquid cultures were centrifuged at 4200 RPM for 10 mins and the pellet resuspended in 10 mL of sterile PBS. This was centrifuged as before and the final pellet was resuspended in 5 mL TSB. Bacterial cultures were normalised in TSB to an optical density of 0.6 nm, making a staring inoculum of 4-

5 x10^6 CFU/mL. A 1:100 dilution was performed by adding 2 µl aliquots of normalised cells to 198 µl of TSB supplemented with 1 % D-(+)-glucose (TSBg) in each of the corresponding wells of a microtitre plate, TSBg was used as this has been found to promote the formation of *S. aureus* biofilms [223,224]. Two hundred microlitres of TSBg was added to a set of wells as blank controls. Microtitre plates with lids were sealed with Parafilm and carefully wrapped in moistened paper towel roll, then placed in a Tupperware box to further prevent moisture loss and contaminants over the incubation time. Plates were incubated at 37 ˚C for 48 h without agitation to allow biofilm formation. These plates were removed after 24 h incubation time, 50 µl of spent media was carefully withdrawn from all test and control wells and replaced with 50 µl of fresh TSBg. Plates were then incubated for a further 24 h at 37 ˚C.

2.6.4 Optimisation of biofilm growth conditions

To determine the optimal growth conditions to promote biofilm formation, a number of approaches were considered and investigated based on previous studies [223,225–227]. Biofilms were formed using the method previously described above, using two preparations of media - TSB and TSBg at concentrations of 100 %, 75 %, 50 % and 25 %. Additionally biofilm formation was assessed using standard microtitre plates and tissue culture microtitre plates (Thermo Scientific Nunc™ Microwell™).

2.6.5 Staining of biofilms

The level of biofilm produced by study isolates was estimated using crystal violet staining and measured by absorbance. Liquid media from biofilm experiment plates was gently removed by inversion onto tissue. Biofilms were washed (twice) by the addition and gentle removal of 200 µl of sterile PBS and then left to air-dry for several hours. Wells were then filled with 200 µl of 0.1 % (w/v) crystal violet (CV) and allowed to sit for 10 minutes at room temperature. Once stained, the CV was poured off and wells were twice rinsed with water to remove residual and unbound CV. The stained biofilms, now visible, were left to air-dry for several hours at room temperature. Stained biofilms were solubilised using 200 µl of 30 % (v/v) glacial acetic acid solution in water deionised water and left for 20 minutes. The contents of each well was resuspended by repeated pipetting to ensure the biofilm was completely solubilised before transfer to a new 96-well microtitre plate. Biofilm mass was estimated spectrophotometrically using a FLUOstar plate reader at absorbance = OD_{590}.

2.6.6 Enumeration of biofilm cells

Colony forming units (CFU) of *S. aureus* cells recovered from established biofilms were enumerated following 48 h incubation (see section 2.6.3). Media was removed from the wells where biofilms were grown and then washed twice with 200 µl of sterile PBS to remove any non-adherent bacteria and residual media. Two hundred microlitres (200 µl) of sterile PBS was added to the selected wells for cell counts and the biofilms were scraped away from the walls and the bottom of the wells using a sterile toothpick. The biofilm suspension was pipetted up and down

five times to homogenise the content within each well. An aliquot of 100 µl of disrupted biofilm suspension was used to make ten-fold serial dilutions in PBS for counting on TSA plates to determine the CFU for each sample. The number of viable cells within each well was calculated using the formula described in section 2.1.2.

2.6.7 Phage lysate preparations

When assessing the effect of phage on biofilms, the number of viable cell recovered from 48 h biofilms were initially obtained in section 2.6.6, phage lysates of a known titre were diluted accordingly in sterile growth media to obtain the desired MOI of 1, a 1/10 dilution was further performed to achieve an MOI 0.1.

2.6.8 Biofilm diminution using phage

The formation of *S. aureus* biofilms was carried out as above (see section 2.6.3), after 48 h the media was poured off and washed twice with 200 µl PBS to remove any non-adhered cells. Biofilms were treated with 200 µl of diluted phage lysate at two different MOIs :- 1 and 0.1 which were added to each corresponding set of wells according to the layout below (Figure 2.5). Two hundred microlitres (200 µl) of TSBg/PBS was used as a negative control and the blank controls were replenished with 200 µl TSBg. Experiments were performed on three separate occasions. Quantification of the remaining biofilm biomass and viable cells following exposure to study phage for 6 and 24 h was performed as described in section 2.6.5 and 2.6.6.

Figure 2.5: Plate layout used for assessing the anti-biofilm properties of phage against mature *S. aureus* biofilms.
Two MOIs were applied to four different bacterial hosts. Four wells were inoculated allowing one well for viable cell counts.

2.6.9 Biofilm visualisation

Biofilms regularly formed on the walls and the bases of the microtitre plate wells. However the small well size made it difficult to visualise the biofilms using microscopy and therefore the base of the wells had to be cut away from the rest of the plate structure. The contents of each well were removed and biofilms were subsequently washed as mentioned in section 2.6.6. Within a fume cupboard, a sterile burning hot scalpel blade held under a Bunsen burner until glowing was used to obtain a clean cut through the polystyrene well, at roughly 0.5 cm up from the base as demonstrated in Figure 2.6. This was attached to the inside of a sterile petri dish using Vaseline.
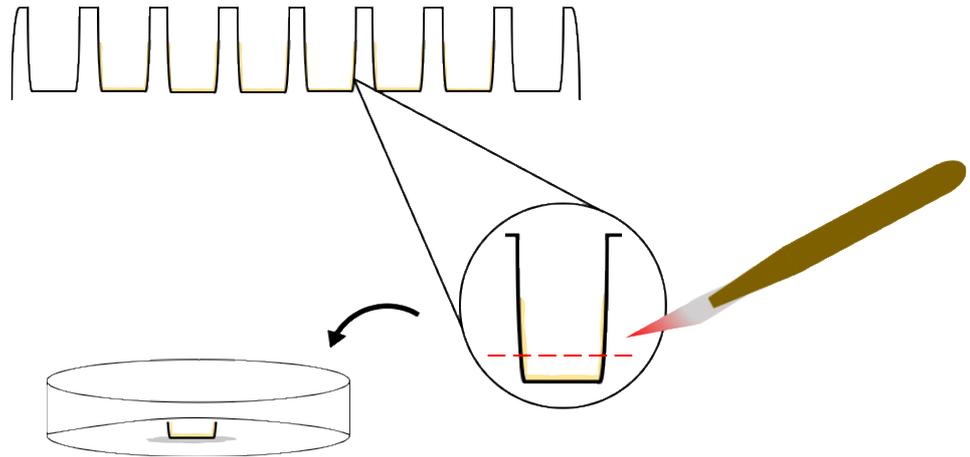
Figure 2.6: Schematic diagram demonstrating the removal of the microtitre plate base containing biofilm

2.7 Confocal laser scanning microscopy

The level of biofilm disruption by phage infection and cell viability was assessed by confocal microscopy. Once placed inside a sterile petri dish, the base of the wells were treated with LIVE/DEAD stain (Live/Dead® *Bac*Light™ Bacterial Viability Kit, Invitrogen Ltd) following the manufacturer's instructions, carefully wrapped in foil and incubated for 15 minutes in the dark at room temperature. The well bases were rinsed several times with PBS to remove any residual stain and 20 mL of water was added in to the petri dish to submerge the well base. Biofilms were observed using a under a confocal scanning laser microscope (Leica TCS SPE 1000 CLSM) equipped with a x40 water dipping lens and examination carried out on Leica Application Suite X (LasX) software package. The LIVE/DEAD was used to stain biofilms following treatment of phage. The staining kit consists of a mixture of SYTO 9 green fluorescent and propidium iodide red fluorescent stains. This allows the

discrimination of live cells (with intact membranes) that stain green and dead cells (compromised membranes) that stain red under CFLM.

2.8 Statistical analysis

Experiments were performed with a minimum of three replicates and these values were used to plot mean ± standard deviation. Statistical analysis was performed using GraphPad Prism Version 7.0 software package, data was analysed as an ordinary one way ANOVA (Analysis of Variance) and Sidak's multiple comparison test to determine significance of results. Results were taken as significantly different by a *p* value of 0.05 unless otherwise stated.

2.9 Phage genome sequencing

2.9.1 Isolation of phage genomic DNA

Fresh phage lysates (10^7 - 10^9 pfu/mL) were prepared on their respective host using the flood plate method (see section 2.3.3). Phage genomic DNA was obtained by a phenol:chloroform isoamyl alcohol (25:24:1 [v/v]) extraction method [228]. For each phage, 1.5 mL of phage lysate was centrifuged at 13,000 rpm for 10 minutes at 4 °C, 1 mL of phage supernatant was transferred into a fresh microfuge tube and treated with DNase I (10 µl of 1 mg/ml DNase I) and RNase A (4 µl of 12.5 mg/ml RNase A) to remove host bacterial DNA and RNA respectively. An equal volume of phenol (pH 10) was added to each tube, vortexed for 30 s and centrifuged at 13,000 rpm for 10 min at 4 °C. Depending on the quality of the phage lysate, two visible layers separated by a white protein interphase should be present within the

tube (Figure 2.7). The aqueous layer on top containing DNA was carefully extracted with a pipette leaving a small volume left over to prevent taking up the lower organic phase or protein interphase. The contents of the pipette was transferred to a fresh microfuge tube and an equal volume of phenol:chloroform (1:1) was added, vortexed for 30 s and centrifuged at 13,000 rpm for a further 10 min at 4 °C. Again, the aqueous layer was collected and transferred to a fresh microfuge tube, an equal volume of phenol:chloroform:isoamylalcohol (25:24:1) was introduced to each tube, vortexed for 30 s and centrifuged at 13,000 rpm for 20 min at 4 °C.
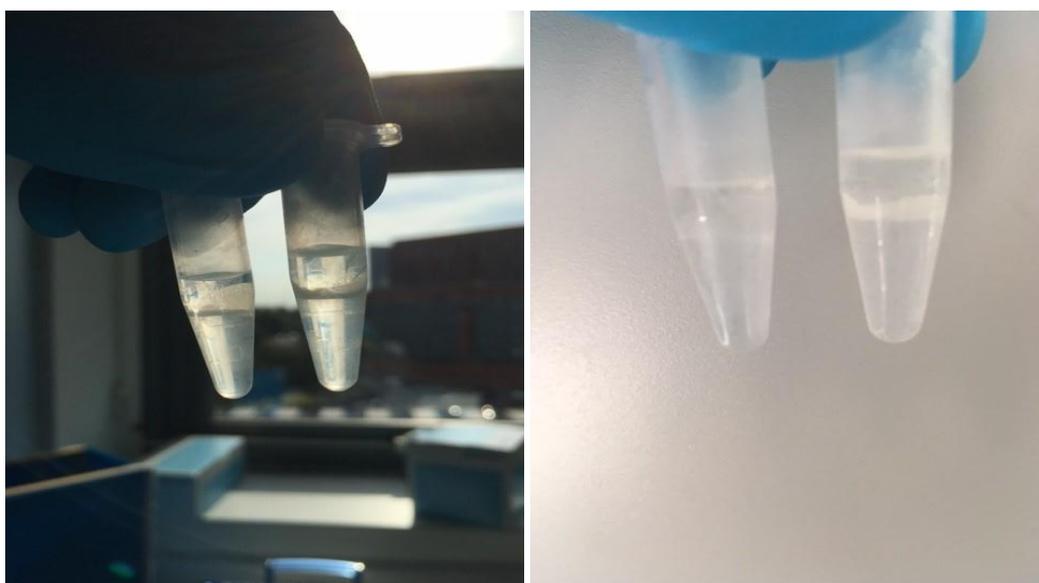
Figure 2.7: Extraction of phage DNA using phenol:chloroform.

The aqueous phase (top) contains most of the DNA, a white layer of denatured proteins partitions the lower organic phase of mostly RNA and lipids in phenol.

2.9.2 Precipitation of phage DNA

The aqueous layer was transferred to a fresh microfuge tube and mixed with two volumes of ice-cold absolute ethanol and 1/10 volume freshly prepared 7.5M ammonium acetate, then allowed to precipitate at -20 °C overnight. The addition

of ammonium acetate was used to remove DNA associated proteins and reduce carryover of polyphenolics, if present, they may affect downstream enzymatic reactions during library preparation and sequencing issues. The following day, precipitated DNA (visible as white cotton strands) was centrifuged at 13,000 rpm for 20 min at 4 °C before absolute ethanol was pipetted off. The DNA pellet was washed twice with 1 mL 70 % ethanol (v/v), centrifuged at 13,000 rpm for 20 min at 4 °C, ethanol was pipetted off and the pellet was left to air dry to allow residual ethanol to evaporate. The pellet was eluted in 100 µl nuclease-free water and incubated at 37 °C to help dissolve DNA.

2.9.3 Quantification of DNA

The purity of each DNA sample was assessed using a Nanodrop One spectrophotometer (Thermo Scientific). The absorbance ratio at 260 nm and 280 nm is generally used as a measure of purity for both DNA and RNA sample extractions. For DNA the ratio considered as pure is ~1.8, values lower than this indicates contamination from possible residual chemicals used during extraction including phenol, guanidine or carbohydrate carryover. Yields were quantified using the fluorescent dye SYBR Green assay, SYBR Green I master mix was prepared using 1 x TE (pH 8), a 1:50 dilution for each purified phage DNA sample was performed in TE. 2-fold dilutions of Lambda DNA (10 ng/µl) were used to plot a standard curve. DNA samples were quantified by fluorescence on a FLUOstar plate reader at an excitation of 485 nM and emission of 535 nM. All DNA was diluted and normalised in 100 µl nuclease-free water to roughly 5 ng/µl using a Qubit 3.0 Fluorometer (dsDNA High Sensitivity (HS) assay, Life Technologies), before further

diluting to 0.2 ng/µl required for sequencing. Measurements were determined using 2 µl of sample and measured in triplicate to obtain an average value.

2.10 Whole-genome sequencing

The NextSeq 500 was Illumina's first benchtop sequencing platform that provides middle and high throughput options in combination with highly complementary microarray scanning, allowing to successfully sequence exomes, transcriptomes and whole genomes. The NextSeq 500 utilises a 2-channel sequencing technology that allows for a significant reduction in cycles duration and overall processing time for each run. NextSeq 500 employs Illumina's highly accurate and reliable sequencing by synthesis (SBS) technology that uses fluorescently-labelled deoxynucleotides triphosphates (dNTP) to sequence millions of clonal clusters. First DNA libraries are created for each sample as mentioned below, DNA fragments are tagged with unique Illumina adapters (oligonucleotides with a known sequence). Additionally, motifs are introduced to those adapters, consisting of sequencing binding sites and regions complementary to the oligonucleotides bound to the flow cell surface. PCR amplification is performed to clonally amplify the library. This adapter technique allows the ability to pool multiple samples in to a single run rather than a single library.

Strands of DNA are amplified multiple times directly on to the surface of the flow cell to generate millions of clonal clusters through bridge amplification. Sequencing is performed on the forward strands and starts when the first single labelled dNTP (or fluorophore) is added to the nucleic acid chain.
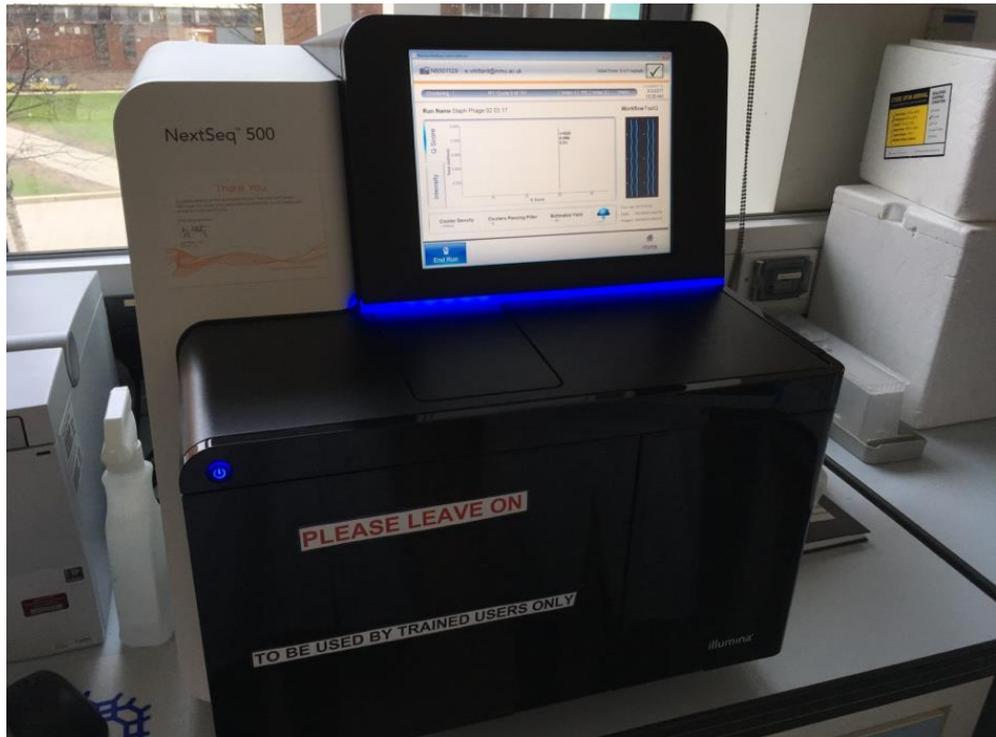
Figure 2.8: NextSeq 500 sequencing platform at Manchester Metropolitan University, UK.

The NextSeq 500, uses a two-colour chemistry to determine all four base calls, by taking only red and green images, base C is labelled with red fluorophore, T is labelled with green fluorophore, when both green and red images are received it is flagged as A, while no emission is flagged as the unlabelled base G base Figure 2.9). Nucleotides are pushed through the flow cell lanes and single dNTPs are allowed to anneal to the these clusters based on the template, once annealed the clusters are excited by a light source, and each dNTP emits a fluorescent signal that is measured and then the fluorophore is cleaved to make way for the following dNTP. During processing, the raw reads from pooled samples are separated based on the unique adapters introduced during the library preparation.
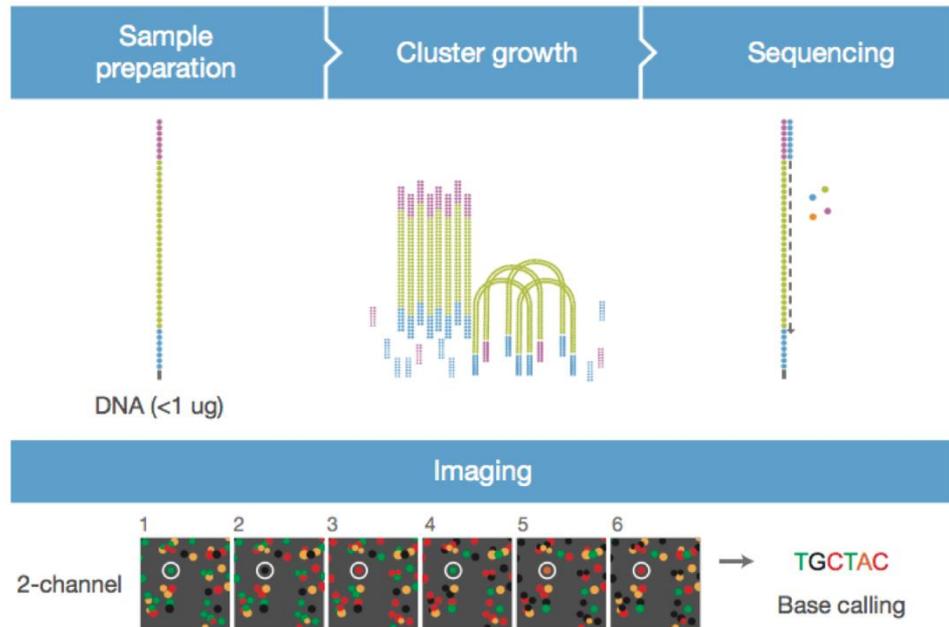
Figure 2.9: Two-channel sequencing by synthesis (SBS) technology employed by the Illumina NextSeq 500 sequencing platform to capture the fluorescently-labelled deoxynucleotide triphosphate (dNTP). Adapted from Illumina.

Libraries of the selected phage DNA samples (input DNA 0.2 ng/μl) were prepared using the Illumina Nextera XT DNA Sample Preparation Kit following manufacturer's instructions. With this kit, the DNA is fragmented and tagged simultaneously by the modified Nextera transposon, quality of each library was assessed and verified using the Agilent Bioanalyzer HS DNA kit and quantified again using a Qubit, the values of each sample was used to calculate the nM for cluster generation using the formula.

Libraries were pooled and diluted to a loading concentration of 1.5pM instead of 1.8pM in attempt to avoid over clustering. Overclustering is most commonly caused by insufficient library clean-up and inaccurate library quantification, which creates a number of negative issues including increases in signal brightness, loss of focus as well as poor template generation, cluster registration and image analysis. High-throughput sequencing of phage DNA (paired-end 2 x 150 High output) was carried out using the Illumina NextSeq500 platform at Manchester Metropolitan University, UK.

2.11 Genome assembly

Next Generation Sequencing can generate a significant amount of sequences data in a single run, the quality control and filtering of this raw data is considered to be one of the most crucial steps during the preliminary stages of data processing. Several kinds of sequencing artefacts have to be processed including poor quality reads, sequencing errors, possible contaminating or overrepresented reads and outlier detection. The identification and subsequent removal (trimming) of these low quality sequence segments will allow for a more successful downstream analysis. Quality of reads were initially checked with FASTQC v0.11.5 [229], lanes were concatenated into a single file for each forward and reverse read. Low quality end reads were trimmed using Sickle v1.33 [230], Trim Galore! v0.4.3 [231] was used to remove the Illumina adapters by trimming 20bp from 5' end and 25bp of 3' end. Sequence data was mapped against a host reference genome sequence using Bowtie2 v2.3.2 to remove host bacterial DNA reads that may affect assembly [232]. SAMtools was used to manipulate each forward and reverse Bowtie2 output

SAM file to a binary BAM file format, the BAM file was then converted to a fastq file, which is much easier for computer programs to read [233]. Assembly of each phage was achieved using SPAdes v3.11 [234]. The pipeline script is shown in Figure 2.10.

All phage assemblies using Spades resulted in a single large contig plus a number of small repeats. The largest contig and their coverage was assessed and visualised using Bandage [235], individual genome assemblies were analysed using Artemis [236] and the largest scaffolds were compared to the similarity of previously sequenced genomes using basic local alignment search tool BLASTn [237].

2.12 Expansion of whole genome phage sequence collection

Based on the similarity between our query sequences and the top hits (closely related genomes) identified using BLASTn, whole genome sequences (FASTA format) of all related phage infecting *S. aureus* were retrieved from GenBank (https://www.ncbi.nlm.nih.gov/nuccore) and the European Nucleotide Archive (ENA) databases in February 2018 to achieve a final collection of 62 phage genomes.

Sickle

sickle pe –sanger –f <inputdir/forward_sequence.fastq> –r

<inputdir/reverse_sequence.fastq>

-o <outdir/forward_sickled_file.fastq> -s <outdir/reverse_sickled_file.fastq> -p singles.fastq

Trim Galore!

perl trim_galore --nextera --length 80 --clip_R1 20 --clip_R2 25 --paired

<inputdir/forward_sickled_file.fastq>, <inputdir/reverse_sickled_file.fastq> -o <outdir>

Bowtie2 build index reference genome database

bowtie2-build <inputdir/TM300_reference_genome.fasta> TM300

Bowtie2

bowtie2 -x TM300 -1 <inputdir/forward_sickled_trimmed.fastq> -2

<inputdir/reverse_sickled_file.fastq> -S <outdir/mapped.sam>

SAMtools

samtools view -bS <inputdir/mapped.sam> > <outdir/mapped.bam>

Samtools bam2fq <inputdir/mapped.bam> > <outdir/mapped.fastq>

SPAdes

spades.py --careful --cov-cutoff auto -k 71,91,111 --pe1-1

<inputdir/forward_mapped.fastq> --pe1-2 <inputdir/reverse_mapped.fastq> -o

<outdir/Phage_spadesout>

Figure 2.10: Whole genome assembly script pipeline from raw sequencing reads.

2.13 Phage genome annotation

Annotation of all study phage and downloaded phage genomes was done from scratch using the same protein database as this ensured that each annotation will be the most up to date and will be the same for all phage. Open reading frames (ORFs) were predicted and annotated with the Prokka v1.12 software tool [238] by searching against the Phaster viral protein database (http://phaster.ca/databases) using default parameters. The script used was as follows:-

```
Prokka

Build database

prokka -- setupdb /location/of/database


Prokka search

prokka –outdir <outdir/Phage_prokkaout> --proteins <inputdir/protein_database.db --locustag <Phage> --prefix <Phage> --genus Caudovirales <inputdir/phage_assembly.fasta>
```

Figure 2.11: Prokka script used to annotate all phage genomes in this study.

The genomes were visualised using Artemis [236] to locate target sequence features and protein sequences. Translated ORFs that were initially annotated as hypothetical proteins were characterised and compared with known proteins using BLASTp, HHpred and HMMER providing further insight in to predicted protein function [239,240].

2.14 Comparative genomics

A neighbour-joining tree was constructed using mash distances using Mashtree [241], the resulting phylogenetic tree was transformed to a cladogram and subsequently converted to Newick format using FigTree v1.4.3 [242]. For pairwise comparisons, Mash sketches were performed for each of the phage assemblies using a default sketch size of 1000, resulting in a single all.msh file using a script (http://s3.climb.ac.uk/ADM_share/run_mash_on_dir.pl) kindly provided by Dr Andy Millard (University of Leicester, UK). To compute all-against-all pairwise distances among the phage genomes, we employed Mash using the all.msh output file [243]. The matrix file was converted by subtracting the similarity value from 1 (e.g 0.2 becomes 0.8 = 80 %), The Mash distance matrix was visualised using the heatmaply package in RStudio [244,245].

Prokka outputs of annotated genomes were used to conduct pan-genome analyses to determine the core and accessory genomes of all study phage. Construction and interrogation of the pan-genome was achieved using Roary with the high speed multiple sequence alignment program (MAFFT) option using default minimum protein BLAST identity at 95 % [246]. Visualisation of pan-genome for all phage and each subsequent cluster was achieved with the interactive visualisation tool Phandango, by using the Roary gene presence/absence output table and the Mash distance tree in Newick format [247]. Genome comparisons were performed using BLAST ring image generator (BRIG) and aligned using Easyfig [248,249].

Predicted proteins were further compared to known proteins using HH-suite [250], for this, protein FASTA output files of the multiple translated CDS sequences for each phage sequence were split into individual FASTA files containing a single sequence using splitfasta.pl tool in the HH-suite software package (https://github.com/soedinglab/hh-suite). Each file was searched against the PDP70 database, outputs produced an individual file for each gene with a summary hit list with a minimum of ten alignments with an E-value cutoff in result alignment of 0.00001, a maximum E-value in summary and alignment list of 0.00001 and maximum coverage with master sequence of 80 %. The scripts used were as follows:-

Mashtree

mashtree numcpus 12 <inputdir/*.fasta> > <outdir/allmash.dnd>

Mash

Mash dist -t <inputdir/all.msh> <inputdir/all.msh> > <outdir/all_v_all.txt>

Heatmaply

Heatmaply (all_v_all_with_percentage, hclust_method="average", sym="T",

main="Staphlococcal_Phage", branches_lwd=0.4, k_row=10 , k_col=10,

margins=c(NA,100,120,40), file="Staphylococcal Phage.html")

Roary

roary -f <outdir/> -e -n *.gff

hhsearch

hhsearch -Z 10 -e 0.00001 -E 0.00001 -cov 80 -i <inputdir/*.faa> -d <inputdir/hh-

suite/Databases/pdb70>

Figure 2.12: Pipeline script used to perform comparative genomics of phage genomes.

Chapter 3

Isolation of novel lytic phage infecting *S. aureus*

3.1 Introduction

Staphylococcus aureus is a global cause of hospital- and community-associated infections that range in severity from simple skin and soft tissue infections to life threatening infections including sepsis, endocarditis and necrotising pneumonia [251]. Its versatility has enabled it to adapt to the selective pressures of antibiotics where it serves as a reservoir for antibiotic-resistant genes within healthcare settings. With resistance to all β-lactam antibiotics, methicillin-resistant *S. aureus* (MRSA) infections are often difficult to treat presenting a major healthcare burden. Invasive MRSA infections are usually treated with the glycopeptide antibiotic vancomycin, which requires intravenous infusion but this is a nephrotoxic agent requiring enhanced patient monitoring [252,253]. Resistance to vancomycin is rare [254], but MRSA with reduced susceptibility (Vancomycin Intermediate *S. aureus* – VISA), are much more common [255–257].

Alternatives to vancomycin exist including linezolid and daptomycin, of which, both have received regulatory approval for use in treating complicated skin and soft-tissue infections [85,258]. However, resistance has emerged to all antibiotics used to treat *S. aureus* and already daptomycin-nonsusceptible phenotypes have been observed in 38 to 83 % of VISA isolates [259]. There is an acute need for new drugs or regimens to treat infections caused by this species. Alternatives or adjuncts to antibiotic chemotherapy were listed in a 2016 report aimed at prolonging the useful life of antibiotics [260]. In this article bacteriophage (phage) therapy was highlighted as being potentially useful based on its history of safe and effective use in some countries.

Attractive features of obligately lytic phage that would be considered for their use in human medicine would include a short adsorption time with a relatively short generation time inside of its host, in addition to producing a high number of viral progeny (burst size) following lysis [261]. This would ensure rapid killing of target cells and a reduction in the chances of bacteria developing resistance to phage infection [159,262], traits that are highly desirable when considering their use within therapy.

The first phase of this project was the isolation and characterisation of a collection of lytic phage. Although phage are generally very abundant in nature, phage infecting a particular species will not be present in the absence of their bacterial host. Recovery of phage from environmental sources where bacterial and phage numbers are low can be enhanced through enrichment of samples with large numbers of a suitable host species to promote the amplification of phage infecting the species of interest [263,264]. *S. aureus* phage have been isolated from water courses, soils and foodstuffs and directly from animal and human faeces. However the richest sources according to most previous studies are sewage / wastewater treatment sites [261,265,266], although isolation from such sites has proven challenging to others [267]. Staphylococcal phage have been recovered from such samples have displayed extensive host ranges compared to phage infecting other species [261]. Two of the most commonly described *S. aureus* phage groups are the Twort-Like (TL) myoviruses and lytic podoviruses (LP) [112,158,205,211,268–270] which have been used to successfully treat human *S. aureus* infections [271,272]. Several of which have demonstrated a polyvalent nature capable of

infecting and killing 70 to 90 % of all clinical *S. aureus* isolates tested against [187,188,269,273,274].

In most published studies that include the isolation of novel phage several techniques are generally used in characterising these new entities. Phage fitness studies help determine the overall growth characteristics of each individual phage, including adsorption rate, generation time, burst size, in addition to understanding their stability under different storage conditions and temperatures. Electron microscopy can reveal morphological features to help better characterise the phage and is used as a major criterion for phage classification [275]. Additionally, it is considered that one of the most important criterion when deliberating the phage therapeutic potential, is its host range, and specificity against test strains. If the host range is too specific this may limit its therapeutic potential, conversely, phage exhibiting a broad host range (polyvalence), capable of infecting, a large number of bacterial hosts, may not lack virulence potential, i.e. high burst size, latent period and required phage dose [276]. The phage can also be studied in liquid culture to examine their *in vitro* kill dynamics and as well as the emergence of phage resistant mutants [262,264,277]. These studies will help determine whether the phage is capable of killing the bacterial host through infection and lysis by phage proteins, and not caused through the 'lysis from without' phenomenon. An event when phage are able to kill bacteria without intracellular replication, as the bacteria is lysed induced by a multiplicity of phage adsorption, thus puncturing the bacterial membrane [278],

This chapter focuses on the isolation and characterisation of novel lytic *S. aureus* phage and examines their efficacy against a large panel of methicillin - sensitive (MSSA) and – resistant (MRSA) isolates.

3.2 Methods

3.2.1 Sample site

Environmental samples were collected from several locations around the north west of England. Initially, animal samples were collected from Fields Farm in Castleton, in the heart of the Peak District National park 28 miles southeast of Manchester and 16 miles west of Sheffield. Samples were also collected from Chelford livestock market, 19 miles south of Manchester (Figure 3.1).
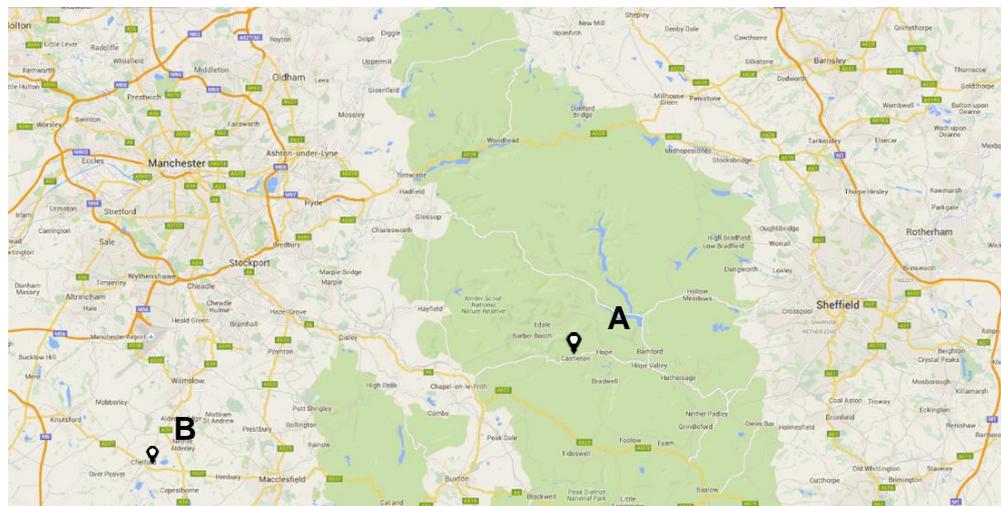
Figure 3.1: Animal sampling site locations.

A) Fields Farm Castleton(Derbyshire, UK), and B) Chelford Cattle Market (Cheshire, UK).
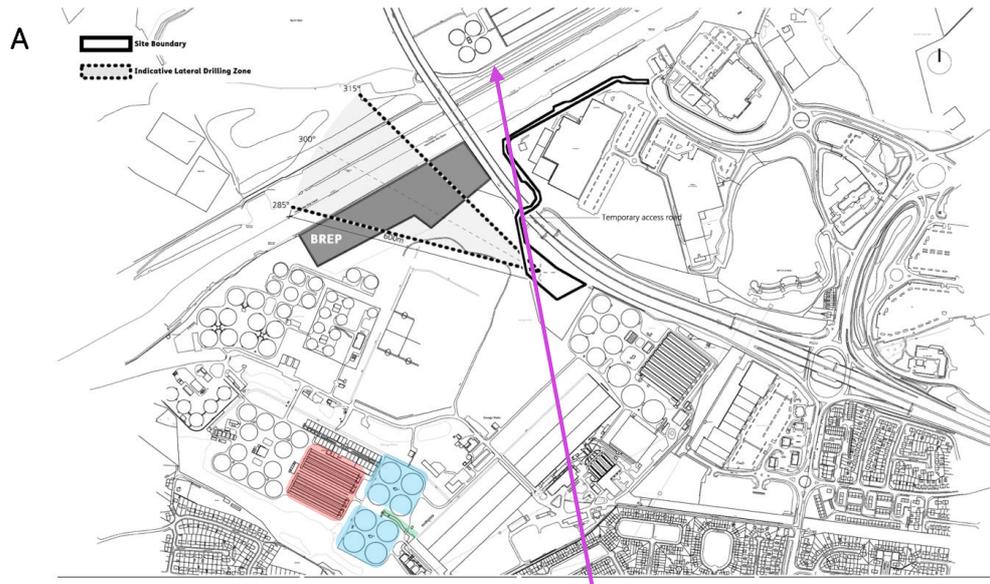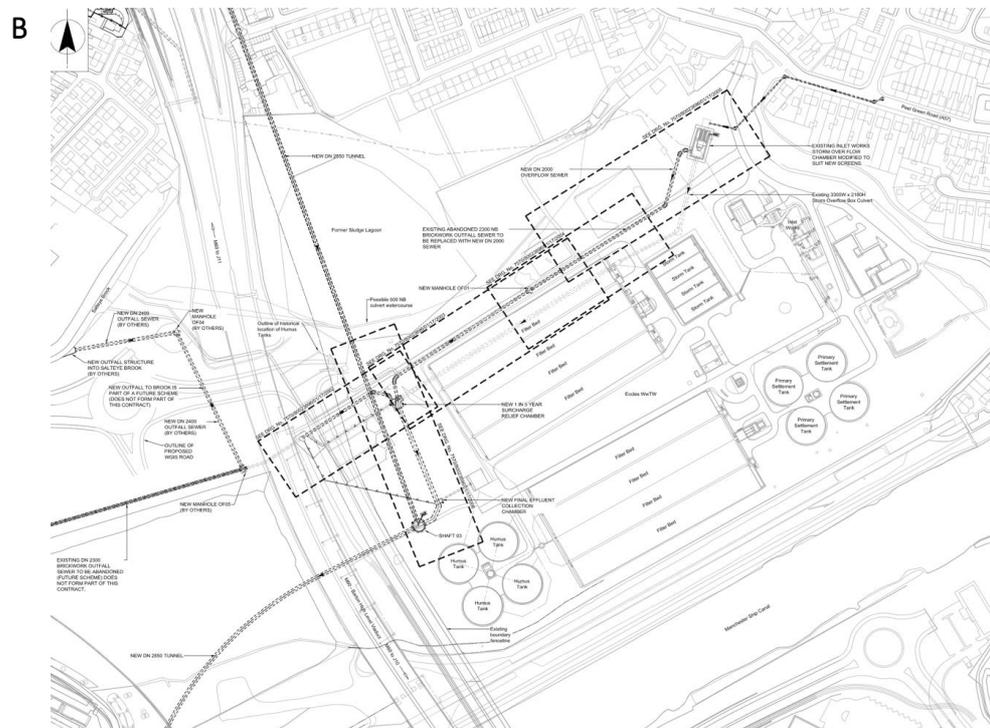
Figure 3.2: Wastewater treatment works sample sites.

A) Davyhulme Wastewater treatment works (WwTW), inlet (green), primary settlement tanks (blue) and ASP tanks (red). The location of Eccles WwTW and its proximity to Davyhulme has been marked. B) Eccles WwTW site layout.

Davyhulme wastewater treatment works (WwTW) is United Utilities largest WwTW serving 1.5 million people for the Greater Manchester conurbation, where it also processes the sludge imports from other WwTWs including Eccles WwTW.

Davyhulme consists of two parallel aerated sludge process tanks (ASP1 and ASP2) made up of 8 separated lanes each, directly downstream of the inlet Figure 3.2A. Eccles WwTW is located just up the canal from Davyhulme WwTW and serves a population equivalent of ~117,000 as of 2018. The primary treatment at Eccles is made up of four primary settlement and sludge tanks with eight biological filter chambers although no sludge treatment on-site Figure 3.2B.

3.2.2 Bacterial Strains

A modified *Staphylococcus carnosus* strain, TM300 kindly provided by Dr Guoqing Xia (University of Manchester, UK) was used to increase the chances of isolating novel lytic bacteriophage from environmental samples, as well as purify and propagate. The antibiotic-resistance marker carried by the TM300 allows all enrichment media to be supplemented with chloramphenicol (10 µg/mL) that would allow the selective growth of TM300. Additionally, an alternate host *S. aureus* D329 was also used for when sample enrichments didn't produce plaques on TM300.

The bacterial strains used in this study to evaluate the lytic ability of the phage collection are listed in Appendix A. The panel is made up of 185 major methicillin – sensitive (MSSA) and – resistant (MRSA) clones from human and animal sources, consisting of 56 different Sequence Types (ST) based on multilocus sequence typing (MLST) see section 1.3.3.1.

3.3 Results

3.3.1 Isolation of staphylococcal phage using samples enrichments

Samples were originally collected from four sources, Castleton, Chelford, Eccles and Davyhulme wastewater treatment works (WwTW) sporadically over a period of a year. However, animal samples were collected until the closure of Chelford livestock market in 2017, and collections from Eccles were discontinued as no phage plaques were observed following sampling over a period of several months. The activated sludge process (ASP) treatment primarily exploits the use of microorganism to break down and metabolise organic matter, therefore each sample should contain extremely high densities and diverse communities of microorganisms. On average, 12 to 16 wastewater samples were collected from different lanes of the two ASP tanks across Davyhulme WwTW on a weekly basis for several months.

3.3.2 Plaque morphology

Phage that were found present in animal faecal samples produced pinpoint sized plaques that were often difficult to pick and purify, these phage were also inconsistent when attempting to propagate and produced very low titre lysates. The screening of around 150 wastewater samples collected from ASP tanks at Davyhulme WwTW, yielded 46 lytic bacteriophage following centrifugation and sample enrichments with the selected host strains. The plaque morphologies of each phage were circular with smooth edges and generally rather small, some with halos but could not be propagated. Plaques varied in size mostly ranging from less than 0.5 mm to 1.5 mm, with EW36 and EW41 producing the largest ranging from

1 – 2 mm and can be seen in Table 3.1. The phage isolated from Davyhulme were found to produce significantly larger plaque with more defined edges than the phage isolated from animal faeces

However, morphological diversity of plaques were still observed even when phage were subjected to several rounds of plaque purification (Figure 3.3). It was also noted that a number of phage produced different plaque morphologies when determining which host to propagate on.
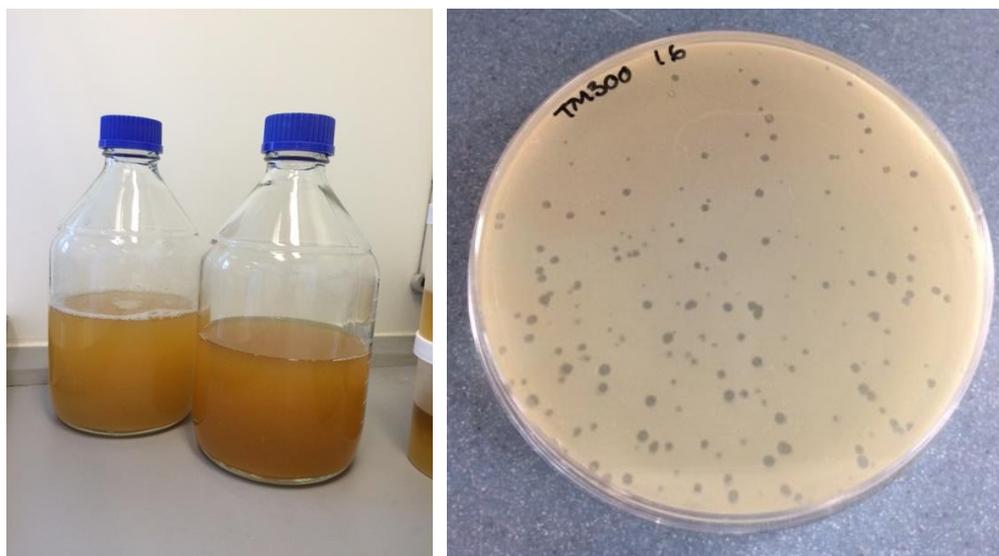
Figure 3.3: Plaques visible on host TM300 using enriched sewage samples collected from Davyhulme wastewater treatment works, UK.

Table 3.1: Isolated staphylococcal phage on their respective host and their plaque size

Phage	Host	Plaque size	Source
EW 1	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 2	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 3	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 4	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 5	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 6	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 7	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 8	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 9	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 10	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 11	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 12	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 13	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 14	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 15	*S. carnosus* TM300	1 - 1.5 mm	Davyhulme WwTW, UK
EW 16	*S. carnosus* TM300	1 - 1.5 mm	Davyhulme WwTW, UK
EW 17	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 18	*S. carnosus* TM300	1 - 1.5 mm	Davyhulme WwTW, UK
EW 19	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 20	*S. aureus* D329	1 - 1.5 mm	Davyhulme WwTW, UK
EW 21	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 22	*S. carnosus* TM300	1 - 1.5 mm	Davyhulme WwTW, UK
EW 23	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 24	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 25	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 26	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 27	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 28	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 29	*S. aureus* D329	0.5 - 1 mm	Davyhulme WwTW, UK
EW 30	*S. aureus* D329	< 0.5 mm	Davyhulme WwTW, UK
EW 31	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 32	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 33	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 34	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 35	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 36	*S. carnosus* TM300	1 - 2 mm	Davyhulme WwTW, UK
EW 37	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 38	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 39	*S. carnosus* TM300	< 0.5 mm	Davyhulme WwTW, UK
EW 40	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 41	*S. aureus* D329	1 - 2 mm	Davyhulme WwTW, UK
EW 42	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 43	*S. carnosus* TM300	0.5 - 1 mm	Davyhulme WwTW, UK
EW 44	*S. aureus* D329	0.5 - 1 mm	Davyhulme WwTW, UK
EW 45	*S. aureus* D329	0.5 - 1 mm	Davyhulme WwTW, UK
EW 46	*S. aureus* D329	1 - 1.5 mm	Davyhulme WwTW, UK

In order to identify the most effective phage against our current panel of bacterial isolates, a collection of 32 staphylococcal phage from previous studies were re-isolated and propagated onto one of the two hosts mentioned above. A collection of seven phage were kindly provided from Dr Guoqing Xia (University of Manchester, UK), the remaining 25 phage were selected from a collection provided by Professor Mark Enright (Manchester Metropolitan University, UK) and Dr Diana Alves (University of Bath, UK). In total, the collection consisted of 78 lytic phage. Concentrated crude lysates were produced from single plaques after three rounds of propagation, with titres ranging from $10^7 - 10^{10}$ pfu/mL, phage lysates were normalised to the same titre of roughly 10^6 pfu/mL for further experiments.

3.3.3 Phage Host range

The host range of each phage was determined by spotting purified phage lysates onto soft agar overlays containing *S. aureus* using 185 genetically diverse *S. aureus* isolates based on MLST. The determination of the host range for each phage was classified based on the degree of clearing on the bacterial lawn following the spot test method (section 2.5, Figure 2.4). Phage varied from each other considerably: 1 - Complete lysis, 2 - Clearing but hazy 3 - Turbid, 4 – Slight disturbance with individual plaques (could possibly be prophage induced from host) as observed in Figure 3.4. Phage that were able to produce 'bullet hole' spots with almost complete clearing but still had a number of phage resistant mutant colonies present within the zone of clearing, were classed as intermediate (Figure 3.5).

96

Overall, the majority of the phage exhibited a broad host range capable of infecting a large proportion of the *S. aureus* collection tested against, 40 out of the 78 phage used in this study were capable of infecting over 90 % panel. From the collection, 15 were capable of disrupting the growth of over 95 % (178 / 185) of the panel. The most effective phage were EW70 and EW71 that were capable of infecting 184 out of the 185 isolates tested against. The one resistant isolate was a MRSA252 phage K mutant isolate, acquired from Dr Diana Alves (University of Bath, UK). EW70 and EW71 capable of completely clearing 53 % (96) and 55 % (101) of the 184 isolates it can infect respectively. When considering the phage ability to cause intermediate and complete lysis to the lawns of the test panel, the phage that was most capable of causing complete lysis was EW74 clearing 59 % (107) of the 183 isolates it can infect, with the lowest consistency of producing phage mutant colonies within the cleared zones.
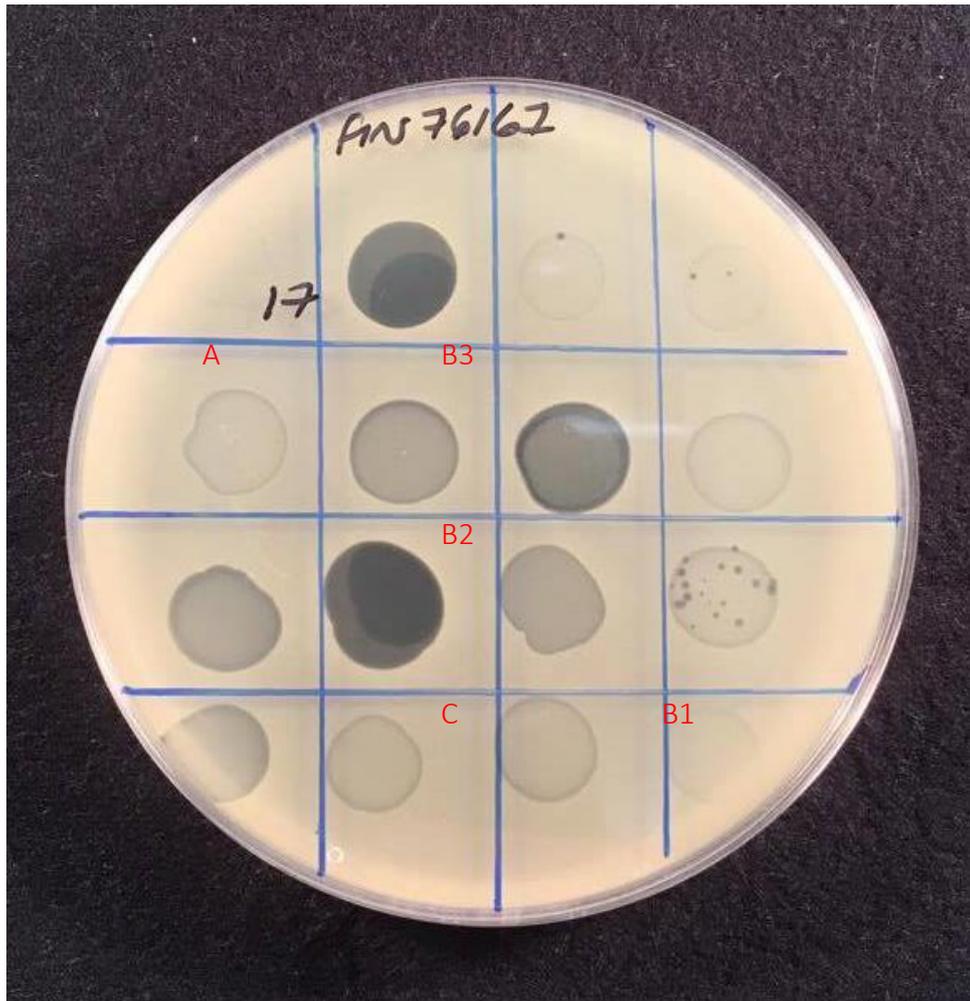
Figure 3.4: Host range assay of 16 EW phages on host FIN 76167.

Plaque formation was scored based on the level of clearing, A) Resistant with no disturbance to lawn, B) Intermediate varied from B1 – Few plaques with slight disturbance to lawn, B2 – Substantial turbidity throughout clear zone, B3 – High degree of clearing of with numerous mutant colonies present, C) Sensitive with complete clearing of bacterial lawn.

Interestingly, both EW16 and EW17 were only able to infect 59 % (109) and 63 % (116) isolates of the 185 isolates in the panel. But of those isolates that they could infect, they were both able to cause complete clearing of the host lawn for 51 % (59) and 52 % (57) of the 109 and 116 isolates respectively.
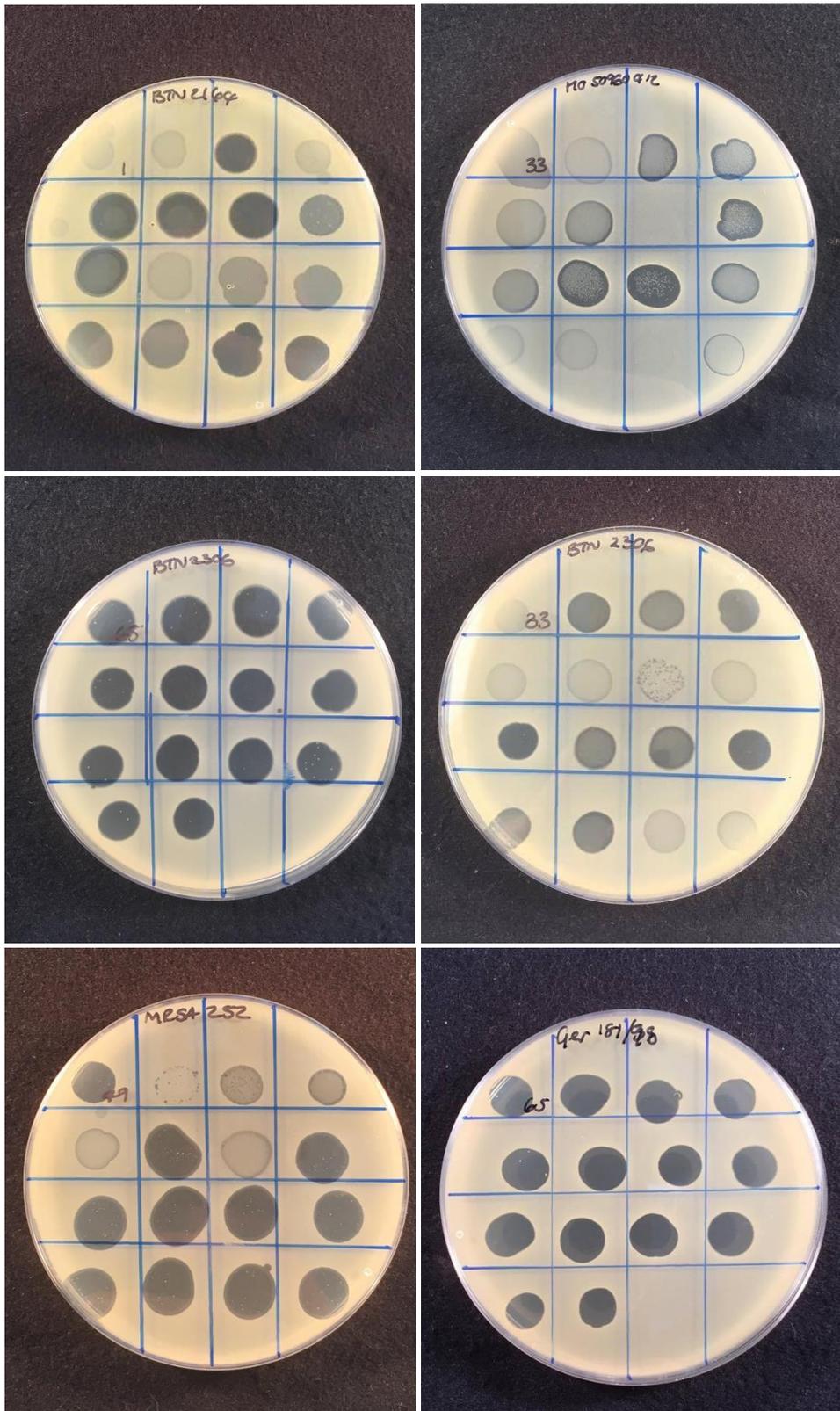
Figure 3.5: A selection of study phages spotted on to the lawns of a number of genetically diverse *S. aureus* isolates

Each host range was performed in triplicate, a number of phage were found to create complete clearing of the bacterial lawn on all three plates with one of the replicates producing phage resistant mutant colonies. In this case they were classed as sensitive, whereas if two of the three plates had mutants then it would be classed as intermediate by majority.

Table 3.2: Percentage coverage of EW phage against 185 *S. aureus* isolates.

The phage host range is scored as Resistant, Intermediate and Susceptible based on the level of clearing on the host lawn. Percentage coverage is the cumulative number of isolates that displayed intermediate and sensitive susceptibility to phage.

Phage	Isolates Resistant	Isolates Intermediate	Isolates Sensitive	Coverage	Phage	Isolates Resistant	Isolates Intermediate	Isolates Sensitive	Coverage
EW1	80	104	1	56.76%	EW40	17	158	10	90.81%
EW2	80	104	1	56.76%	EW41	6	124	54	96.76%
EW3	27	149	9	85.41%	EW42	6	163	16	96.76%
EW4	59	124	2	68.11%	EW43	11	159	15	94.05%
EW5	26	153	6	85.95%	EW44	44	112	29	76.22%
EW6	42	137	6	77.30%	EW45	36	147	2	80.54%
EW7	48	117	20	74.05%	EW46	65	116	4	64.86%
EW8	163	21	1	11.89%	EW47	123	62	0	33.51%
EW9	70	110	5	62.16%	EW48	60	122	3	67.57%
EW10	59	117	9	68.11%	EW49	14	147	24	92.43%
EW11	97	83	5	47.57%	EW50	139	41	5	24.86%
EW12	71	112	2	61.62%	EW51	15	151	19	91.89%
EW13	82	100	3	55.68%	EW52	6	152	27	96.76%
EW14	80	101	4	56.76%	EW53	27	149	9	85.41%
EW15	7	119	59	96.22%	EW54	16	94	75	91.35%
EW16	69	57	59	62.70%	EW55	76	96	13	58.92%
EW17	76	52	57	58.92%	EW56	19	119	47	89.73%
EW18	3	98	84	98.38%	EW57	16	116	53	91.35%
EW19	70	59	56	62.16%	EW58	14	107	64	92.43%
EW20	71	60	54	61.62%	EW59	13	101	71	92.97%
EW21	18	109	58	90.27%	EW60	12	84	89	93.51%
EW22	26	102	57	85.95%	EW61	13	99	73	92.97%
EW23	37	119	29	80.00%	EW62	12	94	79	93.51%
EW24	30	139	16	83.78%	EW63	12	99	74	93.51%
EW25	46	130	9	75.14%	EW64	11	97	77	94.05%
EW26	7	116	62	96.22%	EW65	13	107	65	92.97%
EW27	5	126	54	97.30%	EW66	12	96	77	93.51%
EW28	19	146	20	89.73%	EW67	11	94	80	94.05%
EW29	6	142	37	96.76%	EW68	11	97	77	94.05%
EW30	16	158	11	91.35%	EW69	10	103	72	94.59%
EW31	19	154	12	89.73%	EW70	1	88	96	99.46%
EW32	54	122	9	70.81%	EW71	1	83	101	99.46%
EW33	19	160	6	89.73%	EW72	5	91	89	97.30%
EW34	36	138	11	80.54%	EW73	13	95	77	92.97%
EW35	4	169	12	97.84%	EW74	2	76	107	98.92%
EW36	7	164	14	96.22%	EW75	8	81	96	95.68%
EW37	14	167	4	92.43%	EW76	10	83	92	94.59%
EW38	36	148	1	80.54%	EW77	13	89	83	92.97%
EW39	142	41	2	23.24%	EW78	10	85	90	94.59%

3.3.4 Phage nomenclature

Phage isolated in this study and the phage provided from previous stocks were initially named EW(n) until the phage morphology and/or family is determined (e.g. EW1). Once sequenced they were deposited on to public databases, the phage were renamed as vB_SauM_EW(n) but still referred to as EW throughout this study. The name denotes a **v**irus of **B**acteria with *S. aureus* **M**yovirus morphology, named EW(n).

3.3.5 Growth kinetics of bacterial hosts

The growth kinetics of two bacterial strains -TM300 and D329 that were used for isolation and propagation and strains MRSA252 and 15981 were used to study phage lytic ability (Figure 3.6). These were also useful when determining the appropriate time point to apply phage during phage growth curve studies, in order to ensure that all host strains used were in log phase when phage were introduced. Additionally, measuring the CFU at different optical densities was essential when calculating the multiplicity of infection (MOI) of phage in order to eliminate the 'lysis from without' phenomenon, the MOI chosen for this study was 0.1. This was also important, as the growth curves indicated the appropriate timings for each bacterial host when it reaches the mid-exponential growth phase. Which is helpful when attempting to determine the phage growth parameters through adsorption time and one step growth curve experiments.

TM300 displayed the longest lag phase (4 h) among the four strains, over 2 hours longer than the others. Once TM300 entered log phase, bacterial concentration increased exponentially within a much shorter period of time compared to the other hosts, peaking around 11 hours, then immediately began to decline as it entered death phase. Whereas with D329, MRSA252 and 15981, the lag phase of these strains was relatively short entering its log phase within 2 hours, followed by a steady growth rate that was still increasing by the 20 h time point. During the assay, it was observed that when 15981 and MRSA252 were both growing, they both have a tendency to form clumps and once big enough would eventually sink to the bottom of each well and/or universal tube, creating large sticky masses of bacteria despite regular shaking intervals throughout the assay which may have
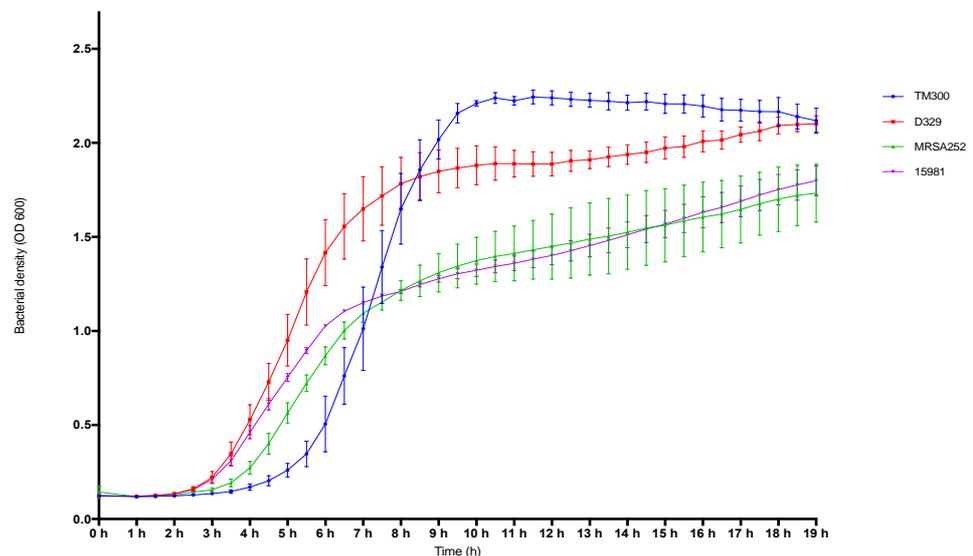
Figure 3.6: Growth curve comparisons of four staphylococcal isolates.

Three *S. aureus* strains D329 (Red), MRSA252 (Green), 15981 (Purple) and one *S. carnosus* TM300 (Blue) in TSB within a microtitre plate. Absorbance readings at 600 nm were taken using a plate reader every 30 minutes for 19 hours whilst shaking at 37 ˚C, three independent experiments were performed in total.

had an effect on the readings. CFU counts and generation time for each host was determined by sampling at different bacterial concentrations of 0.2, 0.3 and 0.4 OD_{540}. CFU counts for each host at each selected bacterial densities remained consistent across all hosts throughout each growth curve assay except for timing of individual growth phases.

3.3.6 Time/kill assay of selected phage against planktonic *S. carnosus* TM300

To elucidate the interactions between the selected phage and its host, we examined its ability to significantly reduce numbers in bacterial broth cultures and study the rate of phage resistant mutant emergence. The ten most effective phage – defined as those with the broadest host range were chosen for further analysis.
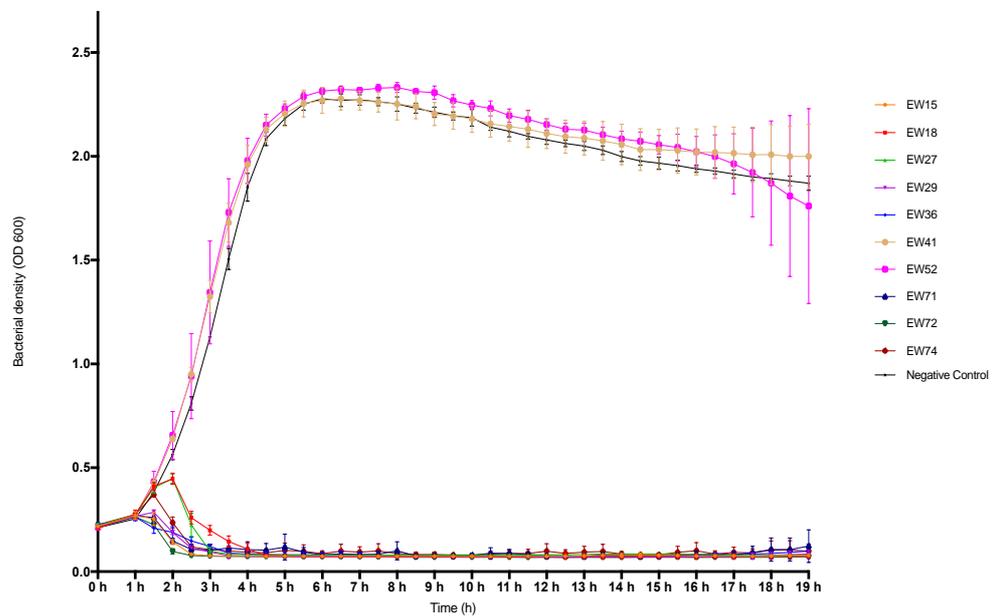
Figure 3.7: Time-kill curve of mid-exponential phase planktonic *S. carnosus* TM300 by selected phage at a multiplicity of infection (MOI) of 0.1.

Absorbance readings at 600 nm were taken using a plate reader every 30 minutes for 19 hours whilst shaking at 37 °C, three independent experiments were performed in total.

Suspensions of these ten phage were tested against strains TM300, D329, MRSA252 and 15981. Phage were introduced to growing cultures to achieve a multiplicity of infection (MOI) of 0.1, and incubated for 19 hours, with each experiment performed in triplicate. Killing curves of the ten phage against TM300 are presented in Figure 3.7. Both EW41 and EW52, propagated on D329 were the only phage ineffective against TM300, the remaining eight phage were successful in reducing the growth of TM300 within 4 hours following introduction and preventing potential phage resistant mutants after 19 hours.

3.3.7 Time/kill assay of selected phage against planktonic *S. aureus* D329

With strain D329, phage EW27 and EW29 initially took an hour longer than other phage before having any effect on the host as seen in Figure 3.8. However, both EW27 and EW29 effectively reduced the growth of D329 after four and six hours respectively, whilst preventing the emergence of phage resistant mutants. As for EW72, it was unsuccessful at depleting bacterial numbers before phage resistant mutants emerged after five hours, although this had an effect on the growth rate of D329 when compared to controls. Interestingly, an increase in bacterial density compared to the control was observed in D329 following addition of EW36. To elucidate the possibility of bacterial contamination within the lysate introduced in to the wells, loops were streaked out on to TSA plates and no bacterial colonies were retrieved following incubation.
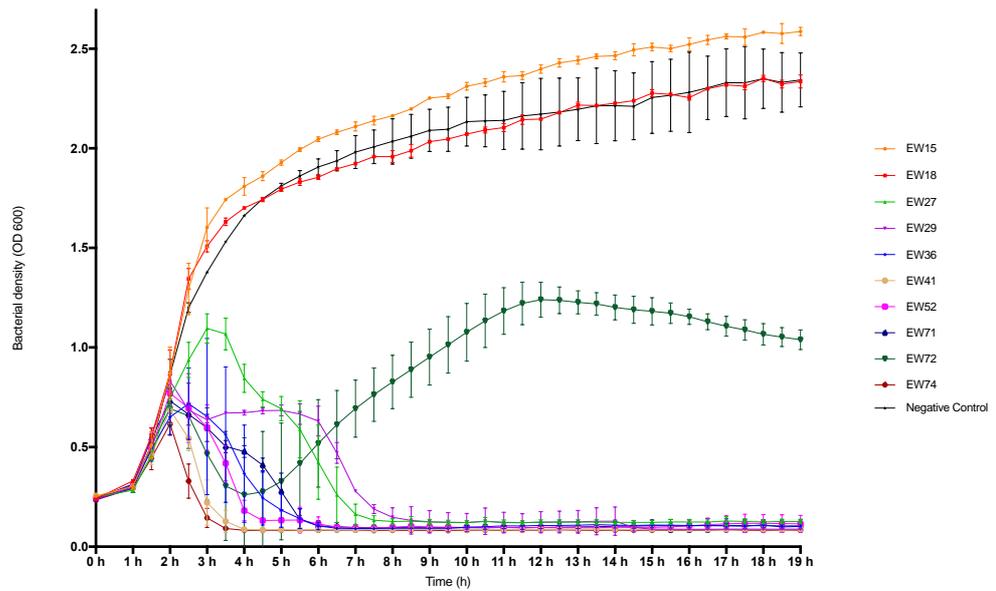
Figure 3.8: Time-kill curve of mid-exponential phase planktonic *S. aureus* D329 by selected phage at a multiplicity of infection (MOI) of 0.1.

Absorbance readings at 600 nm were taken using a plate reader every 30 minutes for 19 hours whilst shaking at 37 ˚C, three independent experiments were performed in total.

Time/kill assay of selected phage against planktonic *S. aureus* MRSA252 and 15981

Individual growth phases appear less clearly defined with MRSA252 when challenged with phage (Figure 3.9), which is also observed with strain 15981 (Figure 3.10) It is clear that a number of phage could not effectively reduce bacterial numbers before resistant mutants emerge. A reduced rate of killing was observed among phage when challenged against other strains compared to their propagating hosts, decreasing bacterial numbers at a much more gradual rate, with some phage such as EW15 taking several hours. Interestingly, phage EW71 and EW74 appeared to have a bacteriostatic effect on the strain 15981 with no change in absorbance observed for ~14 hours before slowly increasing.

As observed in D329, EW36 seemed to have a positive effect on the growth characteristics of MRSA252 when compared to the control as a significant increase

in OD_{600} can be seen. The same trend was also observed with EW72 on the growth of 15981, however, only slightly more than the control. Interestingly, EW52 appeared to have a brief positive effect on the growth rate of MRSA252 for a period of two hours when initially introduced to the wells, followed by a prolonged infection period in which MRSA252 appeared to increase in concentration momentarily, before eventually decreasing in bacterial density below the initial concentration after 15 hours.

For all three *S. aureus* strains used, phage that were able to successfully reduce the optical density and prevent bacterial regrowth were able to achieve it within the initial eight hours relative to the control.
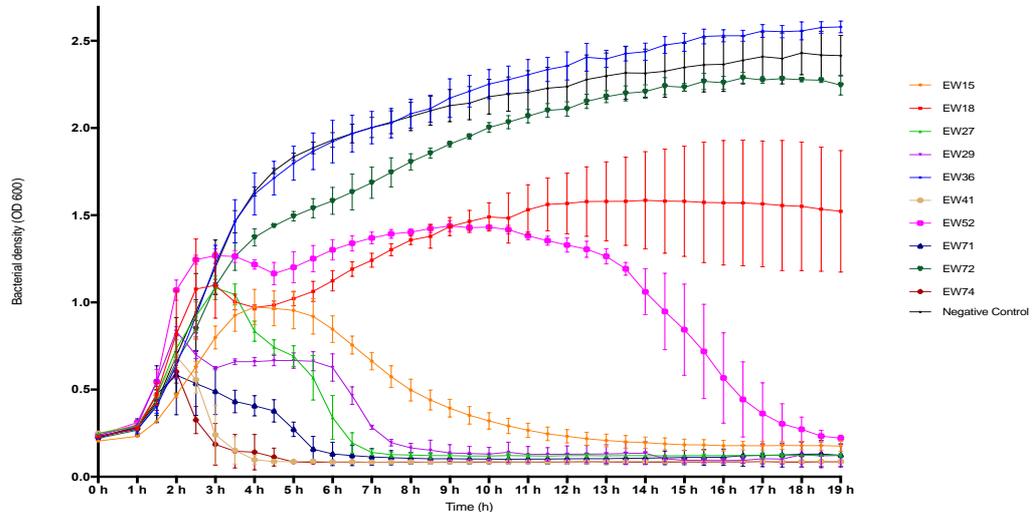
Figure 3.9: Time-kill curve of mid-exponential phase planktonic *S. aureus* MRSA252 by selected phage at a multiplicity of infection (MOI) of 0.1.

Absorbance readings at 600 nm were taken using a plate reader every 30 minutes for 19 hours whilst shaking at 37 ˚C, three independent experiments were performed in total.

Nevertheless, there were a number of phage including EW15, EW27, EW29 and EW52 that exhibited a lower degree of bacteriolytic ability against the selected

hosts, taking up to ~13 hours to have any inhibitory and bactericidal effect on the bacteria. Both EW18 and EW52 experienced resistance from 15981. However EW52 was successful enough to reduce bacterial density eventually, whereas with EW18, although resistance did occur, it had implications on the growth rate of 15981. When challenged with phage, 15981 was found to withstand infection the most, appearing to be the least susceptible to phage infection. Reductions to growth were at a much lower rate compared to other hosts and the appearance of phage-resistance was observed in a number of wells during each of the tests.
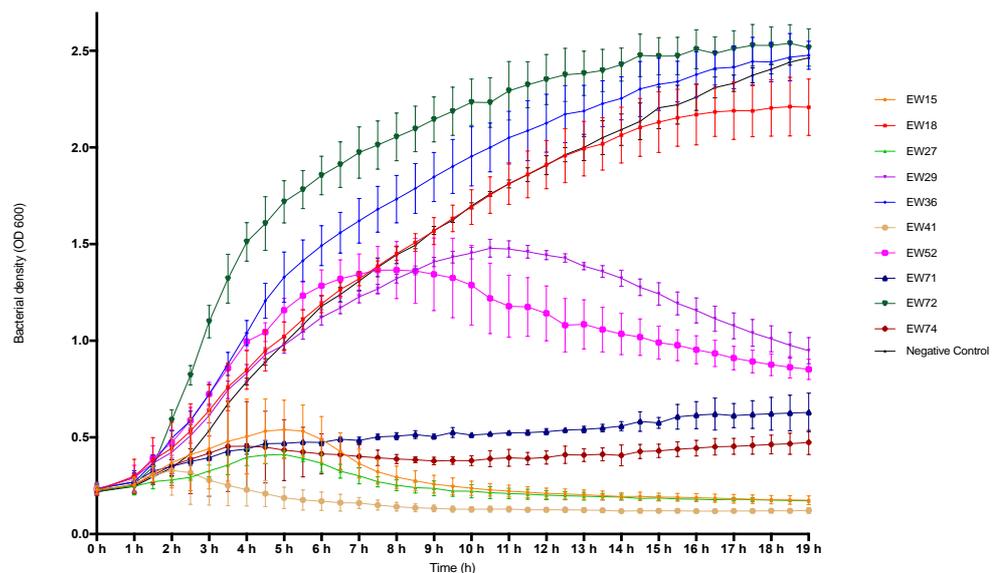
Figure 3.10: Time-kill curve of mid-exponential phase planktonic *S. aureus* 15981 by selected phage at a multiplicity of infection (MOI) of 0.1.

Absorbance readings at 600 nm were taken using a plate reader every 30 minutes for 19 hours whilst shaking at 37 ˚C, three independent experiments were performed in total.

EW41 was the most effective at inhibiting the growth of all three *S. aureus* isolates at the quickest rate as bacterial densities rapidly plummeted within two hours following application. However, EW41 was isolated and propagated on D329 which could be the reason why it had no effect on the growth of *S. carnosus* TM300.

3.3.8 Selection of top four phage

A selection of four phage candidates were chosen for further study on the basis of their lytic characteristics Table 3.3. These exhibited unique growth characteristics as well as a broad host range against all isolates in the panel as described previously, whilst demonstrating a high efficacy in planktonic cultures - effectively and rapidly killing test strains whilst preventing the emergence of phage-resistant mutants. The four phage selected were EW27, EW36, EW41 and EW71. These can theoretically infect 99 % (n = 184) of all 185 bacterial isolates used in this study, of which 69 % (n = 128) are sensitive.

Table 3.3: Top four phage selected following assessment of their lytic ability against the test panel of *S. aureus* isolates.

Phage	Host	Isolates Intermediate	Isolates Sensitive	Total Coverage
EW27	TM300	126	54	97.30 %
EW36	TM300	164	14	97.84 %
EW41	D329	124	54	96.76 %
EW71	D329	83	101	99.46 %

3.4 Discussion

A collection of 78 phage was screened across a large panel of clinically relevant and genetically diverse *S. aureus* isolates including the major clonal groups with high prevalence within the UK and United States. This collection includes 32 uncharacterised phage from previous studies that were kindly provided for this study, and 46 novel lytic bacteriophage that were successfully isolated following enrichments of sewage treatment samples collected from one of the largest wastewater treatment works in the UK.

Phage are widely distributed in the environment and can be found present in various sources including wastewater samples, faecal matter, soil, sea and fresh water [279]. Numerous studies have relied heavily on wastewater and sewage samples as a reliable source to isolate a wide diversity of phage targeting *Escherichia coli*, *Klebsiella* spp. and *S. aureus* [187,280–283]. However, phage presence is largely determined by the existence of their bacterial host within that environment, and although abundant in nature, phage infecting a particular species / strain may be present in very low numbers in a given sample. An enrichment step therefore increases the chances of finding phages of interest from any particular sample where phage numbers may be low.

S. carnosus [284] and *Staphylococcus xylosus* [285]) have been used in previous studies to isolate and propagate staphylococcal phage as these species are avirulent in humans without the virulence factors found in *S. aureus.* This should improve the safety of phage lysates intended for use in healthcare, food and

veterinary applications [285]. The acquisition of a modified *S. carnosus* TM300 was crucial to our study, the plasmid-encoded SaPI enables the expression of a *S. aureus* N-acetyl-D-glucosamine-modified ribitol-phosphate (RboP) repeating units of the wall teichoic acid on to the cell surface of the TM300. These are key phage receptors that may aid with phage adsorption, thus increasing the likelihood of phage isolation. By centrifuging and filtering the samples before enriching, it eliminated the nutrient competition between other bacterial species and allowed our TM300 host to grow unchallenged, resulting in an increased rate of detection and isolation of species-specific phage. The use of double strength TSB during enrichments further maximised the numbers of the growing host by ensuring plenty of nutrients were made available. This approach allowed us to isolate and purify 46 phage from roughly >150 wastewater samples over a period of several months. Achieving a mean hit percentage of 30 %, a significantly higher probability of isolating staphylococcal phage compared to the difficulties reported elsewhere [267]. With the addition of 32 uncharacterised phages kindly provided to us, we assembled a total collection of 78 phage isolates that were successfully propagated on either TM300 or *S. aureus* D329.

The plaques produced when using the TM300 strain as a host were generally small and clear ranging from in size from $< 0.5 - 1.5$ mm, with EW36 and EW41 producing the largest with 'bullet hole' plaques between 1 and 2 mm. One interesting observation throughout the screening of samples, was a number of phage produced plaques with very small halos around them. Although very faint these results were often never replicated during purification effort, with similar

observations also mentioned previously [286]. The presence of these halos suggest the production of soluble enzymes that may have broken down the extracellular polymeric structures from the host strain, which have shown great potential for disrupting biofilms [226,287], as well as with previous phage studies using *Pseudomonas putida* and *Klebsiella pneumoniae* isolates [157,288].

The specificity of each phage was determined using the spot test method against 185 *S. aureus* isolates, consisting up of 56 different Sequence Types (ST). The host range of the phage is largely based on the level of clearing on the bacterial lawn when the phage is spotted on to it, the system used for assessing the success of infection is as mentioned previously [289], where any clearing scored as $1-3$ were classed simply as intermediate. For this experiment, the host range was determined by taking into consideration both intermediate and sensitive results. Our phage collection was successfully able to infect 99 % (184 / 185) of the *S. aureus* isolates tested against, with 40 / 78 phage exhibiting a broad host range (polyvalent), defined as infecting >=90 % of study isolates.

Phage EW70 and EW71 displayed the highest polyvalence and were able to disrupt the lawns of 184 of the 185 *S. aureus* isolates tested, completely clearing the lawns of 96 and 101 isolates respectively. However, the phage that was able to cause complete lysis (59 % - 107 / 183) among the susceptible isolates was EW74, also displaying the least turbidity or frequency of phage mutant colonies present within the cleared zones. As the majority of the phage collection were isolated from environmental sources, it was interesting to see a large proportion of the phage

have lytic capabilities against almost all *S. aureus* isolates from both clinical and animal samples. Polyvalent phage are commonly isolated from natural microbially diverse communities, such as within Activated Sludge Process tanks (wastewater treatment works) where genetic exchange is highly common, and have shown to thrive and reach greater densities within that microcosm [290–292]. The extended host range of our phage collection suggest they have characteristics comparable to most *S. aureus* phage including the most effective Twort-like phage such as ɸ812 [293], Phage K [189] and ISP [294]. All of which have displayed substantial polyvalent activity against study isolates and suitable properties for potential therapeutic application. It must be noted that, the single strain resistant to the entire phage collection was a mutant phage-resistant MRSA 252 isolate, which was specifically chosen for this study in attempt to identify possible phage that could infect this mutant [187].

Using the host range data, the ten most effective phage were selected to evaluate their killing ability against planktonic cultures of three *S. aureus* strains and *S. carnosus* TM300 over a period of time whilst shaking. Killing assays help study the lytic potential of each phage against selected hosts in a more dynamic environment. Providing a much more detailed representation of the bactericidal ability of each individual phage and allows us to observe the rate of killing and mutational events in both bacteria and phage over time. Initially, when performing the killing assays, absorbance readings were taken every 180 seconds. However, when formulating the time/kill graphs, only the measurements every 30 minutes were displayed in order to distinctly identify error bars for each time point.

Throughout these assays, mid-log cultures were exposed to the appropriate number of phages at a multiplicity of infection (MOI) of 0.1, in this environment phages have to adsorb to the host through random collisions rather than spotted on to a static lawn of bacteria [295]. By doing this at an MOI of 0.1, it not only reduces the amount of phage/host collisions but also decreases the probability of the 'lysis from without' phenomenon, a potential limitation from using the spot test assay to determine the phages lytic ability [168]. When applied at low concentrations (MOI 0.1), phage must infect and replicate enough to increase plaque forming units to surpass the rate of replication for the bacterial host. Which would explain why most hosts continued to grow for at least one hour following introduction. In similar studies comparing phage infection at various MOIs [296–298], the greater the MOI the more effective the phage was in the study. Which, presumably was largely due to the increased rate of phage collisions and infections, thus leading to higher densities in viral progeny in a shorter time frame. As well as an increased probability of possible lysis from without as mentioned above. A decline in the optical density (600 nm) throughout the entire experiment was considered to be caused by phage lytic activity for all three *S. aureus* hosts, whereas with *S. carnosus* TM300 a slight decline was expected after 10 hours based from the growth curves. It was observed that the lag phase for TM300 was two hours longer than the three *S. aureus* hosts used, we suspect that this may be due in part to the presence of the modified plasmid it needs to replicate every growth cycle.

In the current study, all of the ten phage selected were able to infect at least one of the four strains tested against, six of which were able to infect and significantly reduce the growth of at least two *S. aureus* isolates. It is interesting to see the variation of results every phage has against each individual isolate. EW41 was found to be the most effective under planktonic conditions, and was able to immediately reduce the bacterial concentration preventing the regrowth of all three *S. aureus* isolates, yet had no effect on the *S. carnosus* isolate. As it was propagated on a *S. aureus* host, the specificity may be limited to just this species. Throughout the killing assays, EW15, EW18, EW36 and EW72 had no effect on the growth of at least one of the three *S. aureus* hosts to some degree. However, phage that were effective were able to prevent the appearance of resistant mutants throughout the duration of the experiment. It was interesting to note that when resistance did occur, the growth rate and fitness for these phage-resistant cells was clearly affected and did not recover to the levels achieved by uninfected controls. This suggests that resistance to phage infection was at the expense of growth capacity for those cells [299,300]. Spontaneous resistance mutation can be associated with changes to the host-specific receptors found on the bacterial cells surface to what the phages adsorb to in order to initiate infection [301]. This could also have an effect on the cells ability to take up nutrients from the environment [302].

When applied at low concentrations (MOI 0.1), phage must infect and replicate enough to increase plaque forming units to surpass the rate of replication for the bacterial host. This would explain why most hosts continued to grow for at least

one hour following introduction. In similar studies comparing phage infection at various MOIs [296–298], the greater the MOI the more effective the phage was in the study. Which, presumably was largely due to the increased rate of phage collisions and infections, thus leading to higher densities in viral progeny in a shorter time frame. As well as a more increased probability of possible lysis from without as mentioned above.

Chapter 4

Phage-mediated dispersal of *S. aureus* biofilms

4.1 Introduction

In the previous chapter, four phages were isolated that have high lytic efficiency and virulence against a panel of *S. aureus* isolates under static and planktonic growth conditions. These phages were further characterised for their anti-biofilm properties against a test panel of clinical *S. aureus* isolates belonging to two major MRSA lineages with global prevalence.

4.1.1 *Staphylococcus aureus* pathology

Staphylococcus aureus is one of the most common causes of biofilm-associated and device-related infections within healthcare. The bacterium presents a burden for healthcare systems and patient safety as it commonly colonises chronic wounds and the surfaces of indwelling medical devices such as intravenous catheters, cardiac pacemakers and joint prostheses [53,54]. Such infections are often difficult to treat using conventional antibiotics, with medical devices having to be replaced much more frequently compared to those infected by other staphylococcal species [55]. A notable characteristic of *S. aureus* biofilms is their ability to produce vast amounts of EPS, composed of teichoic acids, polysaccharide intercellular antigens (PIA), DNA and staphylococcal / host protein components [29]. Their success in biofilms can be linked to their ability to produce an array of proteins known as microbial surface components recognizing adhesive matrix molecules (MSCRAMMS), that can help manipulate the human host defence, aid in adherence and attachment to the surface of cells, tissues and prosthetic devices [18]. Their decreased susceptibility to antimicrobial agents contributes to their persistence in infection.

S. aureus biofilms can be major foci of infection with detachment and dispersal of aggregates through the vascular system leading to colonisation of new niches. This can result in progression of disease from, for example, wound infection to disseminated / systemic disease including endocarditis, necrotising pneumonia, meningitis and septicaemia [56]. *S. aureus* is endemic in hospital environments, and is also a major cause of community-associated infections. These are typically skin and soft tissue MSSA infections but in some locales and population groups community-associated MRSA infections (CA-MRSA) are common in the absence of the risk factors associated with healthcare-associated infections [57].

The local, national and international emergence and transmission of major MRSA clones is a highly dynamic process that can be simply characterised as epidemic waves corresponding to the emergence, spread and eventual decline of clones with novel MSSA genetic background / SCC*mec* combinations [83]. The vast majority of major epidemic MRSA isolates belong to five distinct lineages or clonal complexes (CCs) named after their founder sequence type (ST) CC5, CC8, CC22, CC30, and CC45 [81,82].

ST22 MRSA is the major lineage associated with invasive disease in healthcare within the UK. The rapid dissemination of epidemic MRSA clones belonging to ST22 followed its emergence in the UK in the early 1990s. It then spread across Europe reaching with cases reported in Australia, the Middle- and Far- East [91] where it still remains prevalent [303]. In contrast to ST22, ST36 is frequently isolated in the

UK, however its prevalence has declined more rapidly and is not perceived as successful as ST22 Table 4.1, including other countries as mentioned above [304]. Members of the CC30 lineage include epidemic clone EMRSA-16, isolates of ST36 have are pandemic, found in both hospital and community settings and responsible for many major outbreaks [88].

Table 4.1: Proportion of EMRSA-15 (ST22 - CC22) and EMRSA-16 (ST36 - CC30) isolates collected from 25 clinical laboratories within the UK and Ireland during four study years (2001,2004, 2005 and 2007) [92].

Survey year	EMRSA-15		EMRSA-16		Other strain(s)		Total
	n	proportion of all MRSA (%)	n	proportion of all MRSA (%)	n	proportion of all MRSA (%)	n
2001	78	75.7	22	21.4	3	2.9	103
2003	70	73.7	20	21.1	5	5.3	95
2005	71	81.6	12	13.8	4	4.6	87
2007	76	85.4	8	9.0	5	5.6	89
Total	295		62		17		374

4.1.2 Model biofilm systems

Evaluating the antimicrobial effects of phage in planktonic bacterial cultures may provide useful preliminary data on the potential utility of a phage strain, however their behaviour in experimental biofilm systems may be closer to the *in vivo* situation. A number of studies have demonstrated that bacteria embedded within biofilms exhibit very different behaviour and physiology compared to their planktonic counterparts [305,306] and therefore the effects of any antimicrobial agent should be tested under such conditions. Studies using biofilm models may be a good way to determine the pathogenicity of some species due to the varying levels of biofilms produced from individual isolates [223].

There are several experimental model biofilm platforms that have been established to simulate the *in vivo* environment to help address fundamental questions about biofilm structure and physiology. Data using these models can help provide preliminary empirical data to guide future *in vivo* (animal model) and hopefully, human therapeutic studies.

4.1.3 Lytic phage as biocontrol agents

The experimental use of phage and their derivatives have been demonstrated in many studies as a means of targeting and controlling clinically relevant pathogens responsible for biofilm-associated infections. Their ability of phage to self-replicate at the site of infection, increasing in numbers thus requiring smaller initial doses makes phage an appealing choice as an alternative or complimentary therapeutic agent. Whole phage have been observed to penetrate the glycocalyx reaching deep-lying bacterial cells within the biofilm matrix that were previously protected [307]. Studies have also used phage lytic enzymes that target the key components of the biofilm matrix, causing the biofilm biomass to detach and disperse whilst demonstrating promising results in the control and prevention of biofilms [308,309]. Such enzymes include holin, lysin and depolymerases that target the host cell wall, and thus have the potential to degrade the biofilm matrix as discussed in section 1.4.5.

Phage candidates for human therapeutic use should ideally be obligately lytic and polyvalent - capable of infecting multiple strains of a pathogenic bacterial species. Such phage may have potential as biocontrol agents in *S. aureus* biofilm models.

The most effective and frequently studied are the *S. aureus* phage belonging to the Twort-like myoviruses [187,202,204,294] and lytic podoviruses [270,309,310]. These have good bactericidal activity in both *in vitro* and *in vivo* models, infect multiple strains of *S. aureus* including MRSA isolates as well as coagulase-negative staphylococcal (CONS) species [5].

4.2 Extended methodology

4.2.1 Bacterial strains

S. aureus strain 15981 was used in initial biofilm studies to help optimise growth conditions as this strain is an abundant biofilm producer with excellent adherence properties [198,311,312]. However this strain was not used throughout this project as the aims of this work were to characterise lytic phage activity against isolates of clinically important MRSA lineages. *S. aureus* ST22 and ST36 isolates used in this study are listed in Table 4.2, 14 ST36 isolates were kindly provided by Professor Ross Fitzgerald (The Roslin Institute, University of Edinburgh, UK) to increase the variety of study isolates. Isolates selected for this study were chosen on the basis that the majority of their isolates had their genomes sequenced and were publicly available. This was to allow for potential future genome-wide association studies (GWAS) to identify genetic associations with biofilm formation and phage susceptibility phenotypes.

Table 4.2: List of clinical S. aureus strains with their sequence type (ST) and accession.

Isolate	ST	Accession	Isolate	ST	Accession
1018.07	22	SAMEA957299	M810 08	22	SAMEA957236
1091	22	SAMEA957277	NL011399-5	22	SAMEA957209
370.07	22	SAMEA957217	RH 0600 0061/09	22	SAMEA957269
403.02	22	SAMEA957188	SwedenAO9973	22	SAMEA957194
434.07	22	SAMEA957182	T277 06	22	SAMEA957222
723.07	22	SAMEA957184	T505 30	22	SAMEA957234
729192 April	22	SAMEA957186	W449 36	22	SAMEA957235
921.07	22	SAMEA957210	WW1678/96	22	ERS034629
930.02	22	SAMEA1030324	UK96/32010	36	SAMEA957275
98.4823.X	22	SAMEA957295	NottmA2	36	SAMEA957246
98/10618	22	ERS002153	NottmA	36	SAMEA957262
99ST18131	22	SAMEA957293	MRSA252	36	BX571856.1
AR 0650 784	22	SAMEA957185	H352	36	Not Available
ARI 10	22	SAMEA957259	H119 MRSA	36	SAMEA957179
ARI 11	22	SAMEA957266	FIN75916	36	SAMEA957178
ARI 12	22	SAMEA957203	EMRSA 16	36	SRS003281
ARI 26	22	SAMEA957298	CDC 960758 USA 200	36	SAMN02314232
ARI 29	22	SAMEA957258	BTN 766	36	Not Available
ARI 31	22	SAMEA957255	BTN 2292	36	SAMEA957180
ARI 4	22	SAMEA1030317	BTN 2172	36	SAMEA957297
ARI 5	22	SAMEA957218	BTN 1429	36	SAMEA957183
ARI 7	22	SAMEA957177	98.5806.F[†]	36	SRR453046
BTN 1626	22	SAMEA957294	97.2483.Hb[†]	36	SRR453045
C101	22	SAMN02595338	07.7206.Y[†]	36	SRR453044
C720	22	SAMEA957287	07.6659.K[†]	36	SRR453043
EMRSA15-90	22	SAMN02767598	07.6636.Y[†]	36	SRR453042
F869 56	22	SAMEA957239	07.3481.N[†]	36	SRR453041
H182MRSA	22	SAMEA957195	07.2880.V[†]	36	SRR453040
H431 62	22	SAMEA957230	07.2589.M[†]	36	SRR453039
H65	22	SAMEA957286	07.2496.L[†]	36	SRR453038
H914 91	22	SAMEA957233	07.2449.K[†]	36	SRR453037
HO 5096 0412	22	HE681097	07.1696.F[†]	36	SRR453034
HO 5322 0548 09	22	SAMEA957260	07.1227.Z[†]	36	SRR453033
HO 7230 0407/05	22	SAMEA957292	06.9570.L[†]	36	SRR453032
HO 7374 0468	22	SAMEA957289	03.1791.F[†]	36	SRR453031

[†] Isolates kindly provided by Professor Ross Fitzgerald (Roslin Institute, UK)

122

4.2.2 Phage on biofilm study

For this study, a closed-system biofilm model was selected as the experimental platform. This *in vitro* model system is the most commonly used method employed for studying bacterial biofilms in their simplest form, it was therefore considered to be a good platform to study the bactericidal properties of phage. To help optimise the system, three approaches were employed to investigate the difference in cell attachment and biofilm formation based on: i) Two preparations of liquid media - TSB and TSBg (supplemented with 1 % D-(+)-glucose) ii), Different concentrations of TSB media at 100 %, 75 %, 50 % and 25 % iii) Using standard polystyrene or cell culture microtitre plates.

Viable cells recovered from 48 h biofilms were enumerated in order to calculate the required multiplicity of infection for phage preparations according to each host. Phage lysates were initially prepared and diluted in the selected growth media used to support biofilm formation as above.

4.3 Results

4.3.1 Optimisation of biofilm growth conditions

For this assessment, *S. aureus* 15981 was used due to its ability to form significant amounts of biofilm. Microtitre plates were incubated for 48 hour to allow biofilms to form, wells were washed twice, stained with crystal violet and allowed to dry. Dried biofilms were solubilised in 30 % acetic acid and their densities were measured at an OD_{590} nm. Higher levels of biofilms were present in wells containing TSB supplemented with 1 % D-(+)-glucose (TSBg) at concentrations of 100, 75 and

50 % TSBg when compared to wells with TSB alone. Biofilms grown in 100 % TSBg

produced the highest levels of biofilm biomass with an average optical density of

2.7 (590 nm). There was a significant difference between 100 % TSBg and 100 %

TSB concentrations ($p < 0.01$), however, when 100 % TSBg was compared against

other concentrations, there was no significant difference between 75 % TSBg, 75

% TSB and 50 % TSBg. Additionally, when comparing biofilm densities between

tissue culture and standard microtitre plates, there was no significant difference

between media at the same concentrations. However, significant differences were

observed when 100 % TSBg and 75 % TSBg grown in tissue culture plates against

lower concentrations using standard microtitre plates.
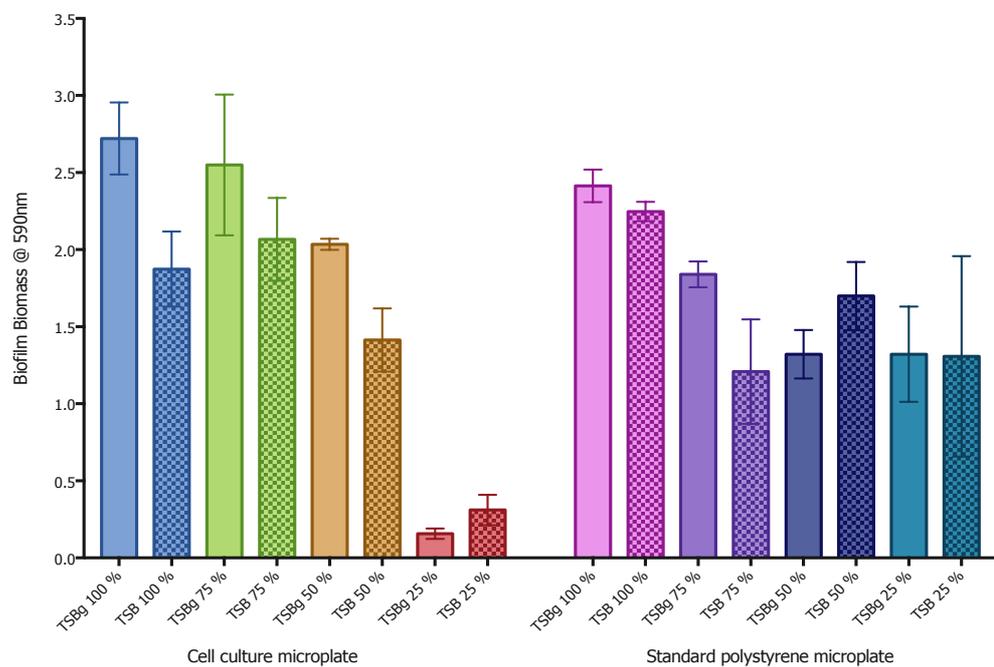
Figure 4.1: Relative biofilm formation of *S. aureus* isolate 15981 after 48 hour incubation within different growth conditions.

Biofilms were assessed when grown in various concentrations of growth medium, with and without the presence of D-(+)-glucose using standard and tissue culture microtitre plates. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

4.3.2 Assessment of *S. aureus* ST22 biofilm production

To identify appropriate candidates for use in future biofilm studies, a collection of 43 *S. aureus* ST22 isolates were studied for their biofilm producing capabilities. It must be noted that, only ST22 isolates from our collection that have had their complete genome sequenced were included in this study. Experiments quantifying biofilm formation by 43 ST22 strains were assessed by crystal violet staining and their absorbance measured (OD_{590}) using a FLUOstar Omega plate reader to determine the most abundant biofilm producing isolates. Biofilm densities of isolates are shown in Figure 4.2. The average optical density across all strains was 0.254. A nominal value of 0.4 OD_{590} or above was consequently used to classify bacteria as strong biofilm producers. From the 43 strains, only seven were considered strong biofilms formers above the cut-off value, with isolates ARI 10 and W449 36 producing the highest levels of biofilm production that were significantly different to all other strains in the collection ($p < 0.001$). Furthermore, there was no significant difference between the nine strains also above the threshold. From this work, four isolates - ARI 10, W449 36, 1018.07 and 370.07 were selected for biofilm experiments using lytic phage.
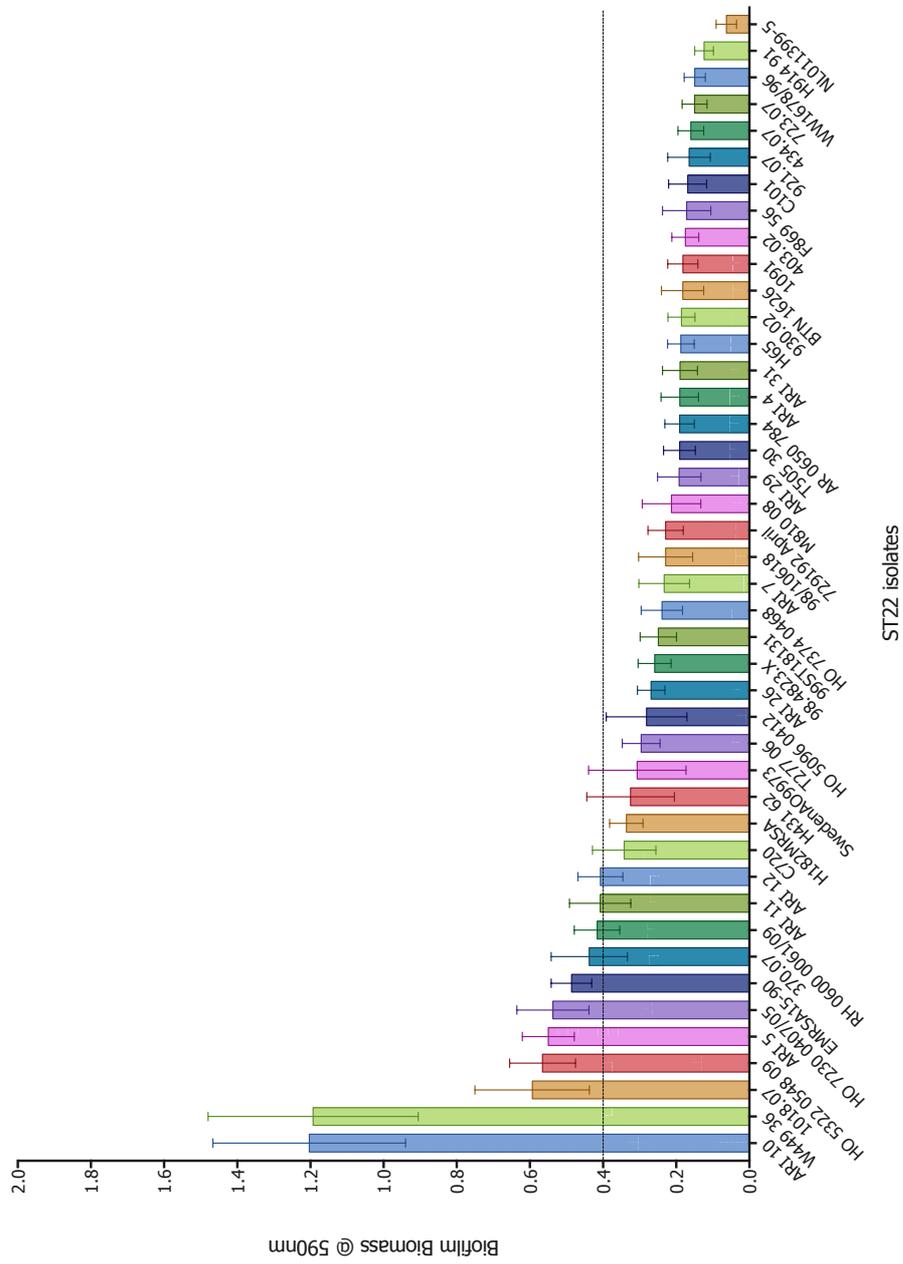
Figure 4.2: Biofilm formation of 43 S. aureus ST22 strains grown in tissue culture microtitre plates over 48 hours at 37 °C.

Biofilms densities produced were assessed following crystal violet staining, dissolved in 30 % acetic acid and measured at an OD590. The cut-off value indicated by the dotted line was used to determine the "strong" and "weak" biofilm producers. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

4.3.3 Assessment of *S. aureus* ST36 biofilm production

A selection of 27 genome sequenced ST36 isolates were available for use in this study. The isolates ability to form biofilms was assessed as above in and can be seen in Figure 4.3. The majority of ST36 strains were capable of producing strong biofilms with an average optical density of 1.93. The level of biofilm formation had higher variance among ST36 strains compared to those produced by ST22 isolates, of the 27 strains within the collection, 25 produced biofilms with an OD of 1.2 or above. The two strongest biofilms formers – isolates 07.1696.F and 06.9570.L along with the two weakest biofilms formers BTN 2172 and 07.2496.L were selected for further study. Both BTN 2172 and 07.2496.L were selected on the basis that they produced biofilms corresponding to OD_{590} values below or equal to 1.2, which were similar to the best ST22 biofilm formers. It was thought that this would allow for better comparisons of phage activity between and within ST22 and ST36 isolates.
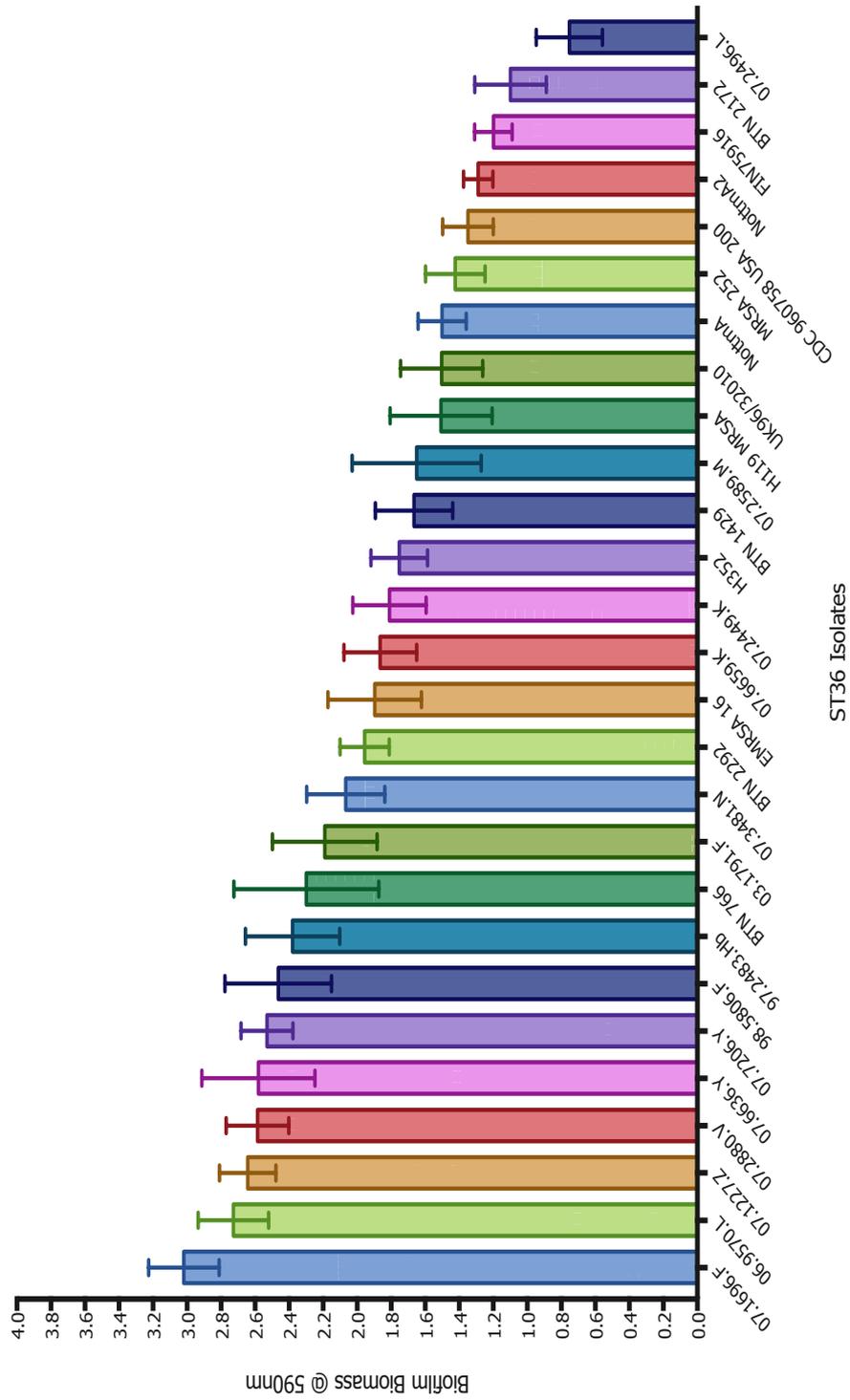
Figure 4.3: Biofilm formation of 27 _S. aureus_ ST36 strains grown in tissue culture microtitre plates over 48 hours at 37 °C.

Biofilms densities produced were assessed following crystal violet staining, dissolved in 30 % acetic acid and measured at an OD590. The cut-off value indicated by the dotted line was used to determine the "strong" and "weak" biofilm producers. Each assay was performed in triplicate, data presented as

4.3.4 Enumeration of viable bacterial cells from 48 hour mature biofilms

In order to quantify the number of viable cells in biofilms, and relate this to crystal violet absorbance, eight *S. aureus* isolates were selected for study. These compromised four ST22 isolates - ARI 10, W449 36, 1018.07, 370.07 and four ST36 isolates - 07.1696.F, 06.9570.L, BTN 2172 and 07.2496.L. In order to obtain quantitative information about the composition of each biofilm, 48 hour mature biofilms were established and viable cell counts were performed as represented in Figure 4.4 These results show that the variation in biofilm biomass between strongest and weakest biofilm formers does not correlate to the number of viable cells present within the biofilms. In terms of colony forming units (CFU/mL) recovered from the 48 h biofilm, isolate 06.9570.L produced the highest viable cells found present within each biofilm ($7.9x10^7$ CFU/mL), whereas the lowest, from isolate 07.2496.L, had the fewest biofilm numbers ($1.8x10^7$ CFU/mL). There was a significant difference between recoverable cells present in the biofilms of isolates 07.1696.F and 06.9570.L ($p < 0.01$) with isolate 06.9570.L producing twice as many cells, however there was no significant difference between the optical density 590 nm of biofilm biomass. Interestingly, when comparing the CFU counts of isolates 07.1696.F and 370.07 there is no significant difference between both, but they are very different as measured by optical density.
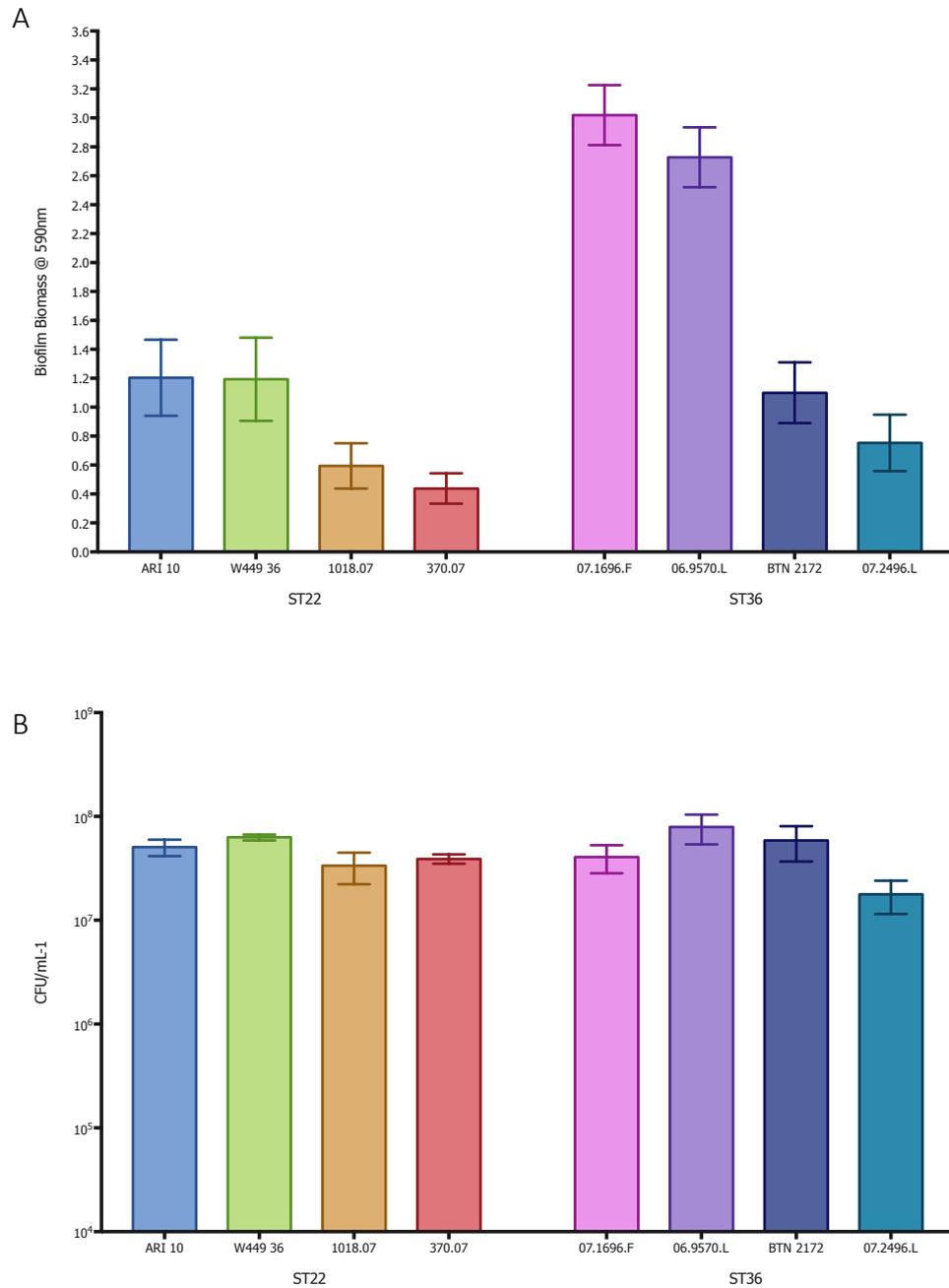
Figure 4.4: Biofilm densities of eight *S. aureus,* including four ST22 and four ST36 strains grown in cell culture-treated microtitre plates over 48 hours.

A – Biofilms initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

ST22 strains displayed a propensity to form moderately adhered biofilms that had significantly lower ODs than the best ST36 biofilm formers, yet they had consistently higher cell counts - similar to those values observed from ST36 isolates. The levels of biofilm biomass developed by each of the ST22 and ST36 isolates after 48 h can be can be observed by eye following crystal violet staining. The biofilms produced by *S. aureus* were found at the air-biofilm interface, although tended to form large aggregates leading to extensive solid-liquid colonisation at the base of the microtitre plate wells as pictured in Figure 4.5.

| ARI 10 | W449 36 | 1018.07 | 370.07 |

| 07.1696.F | 06.9579.L | BTN 2172 | 07.2496.L |
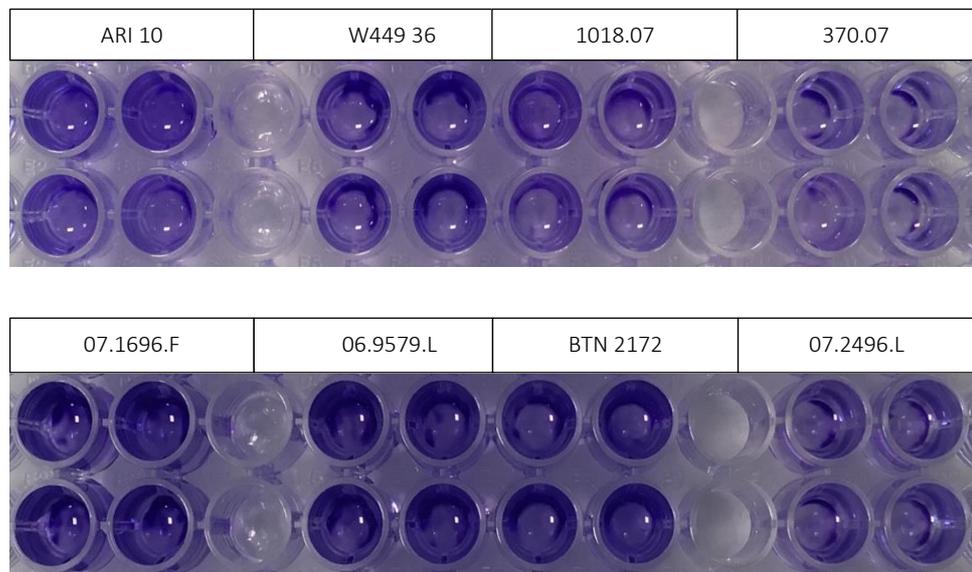
Figure 4.5: Variation in biofilm production by ST22 (A) and ST36 (B) isolates after 48 h incubation in TSB supplemented with 1 % D-(+)-glucose. Biofilms were visualised following staining with 0.1 % crystal violet.

4.3.5 Assessing the anti-biofilm properties of phage EW27 against ST22 and ST36 isolates

Mature (48 hour) biofilms of four ST22 and four ST36 strains were challenged with EW27 lysates at a multiplicity of infection (MOI) of 1 and 0.1 PFU per viable *S. aureus* cell. Viable cell counts and absorbance measurements were taken after 6h and 24h exposure and compared to controls - Figure 4.6 and Figure 4.7. Viable cell counts recovered from each phage-treated biofilm for all ST22 and ST36 isolates except for isolate 07.2496.L, were significantly reduced ($p < 0.001$) following six hour exposure to EW27 when compared to untreated biofilm controls.

However following an initial decrease in CFU/mL after six hours, an increase in bacterial concentration can be seen across all phage treated ST22 isolates after 24 hours, suggesting resistance to phage had occurred within that time. There was no significant difference between CFU counts from wells treated with phage for 6 and 24 hour ($p < 0.05$). When considering the overall biofilm biomass following phage exposure and crystal violet staining, results revealed phage EW27 was highly effective at reducing the biofilms produced by all ST22 and ST36 strains.

For EW27 treated ST36 isolates, biofilm biomass significantly increased in isolates 07.1696.F, 06.9570.L and BTN 2172 using both MOI 1 and 0.1 ($p < 0.05$), despite a minor reduction in bacterial numbers after 24 hour treatments compared to 6 hours. Following treatment of EW27 after both timepoints, EW27 at a MOI of 0.1 proved to be the most effective at both reducing bacterial cells and biofilm biomass for almost all ST22 and ST36 strains.
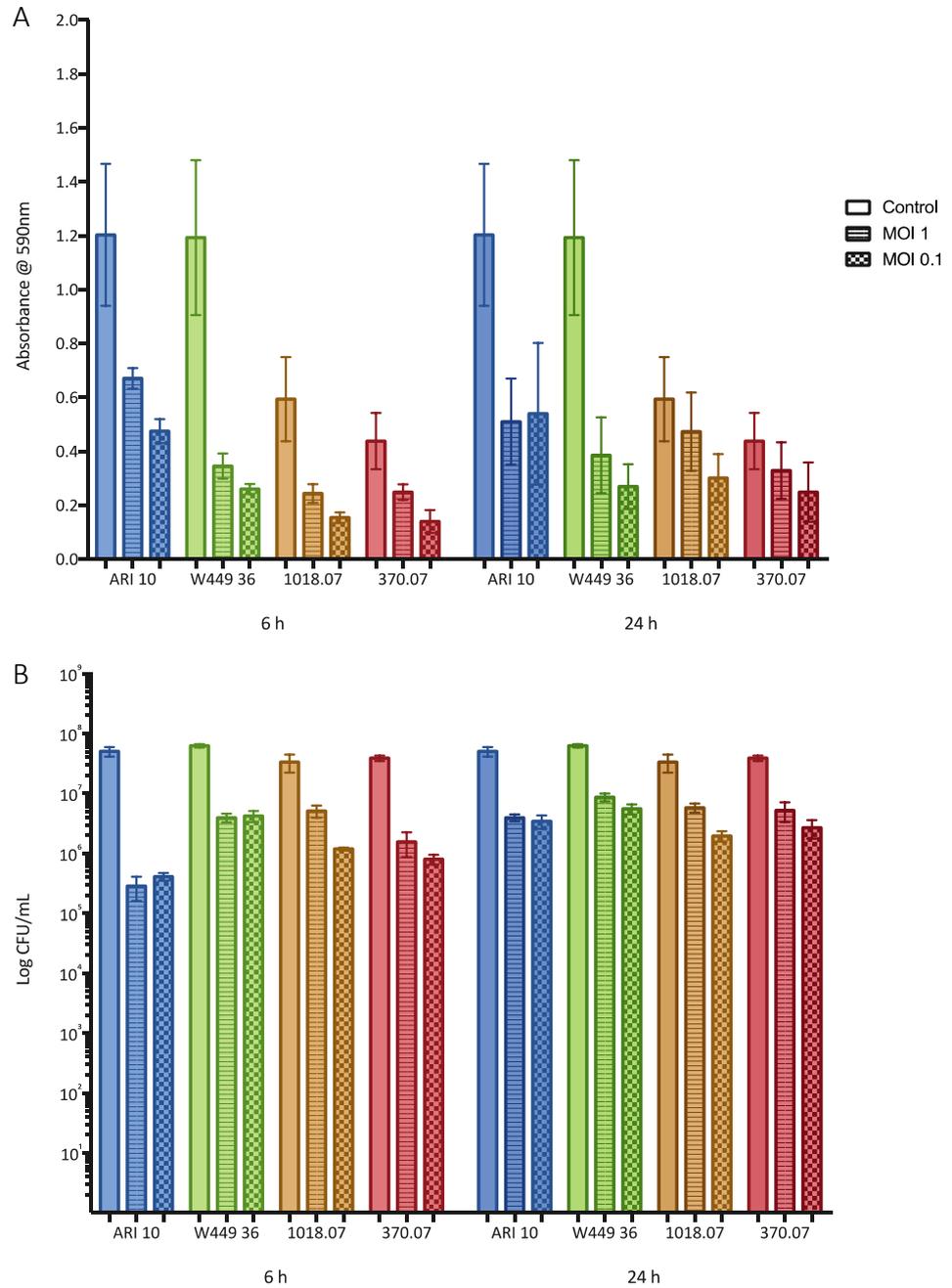
Figure 4.6: Effect of phage EW27 on mature biofilms of *S. aureus* ST22 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).
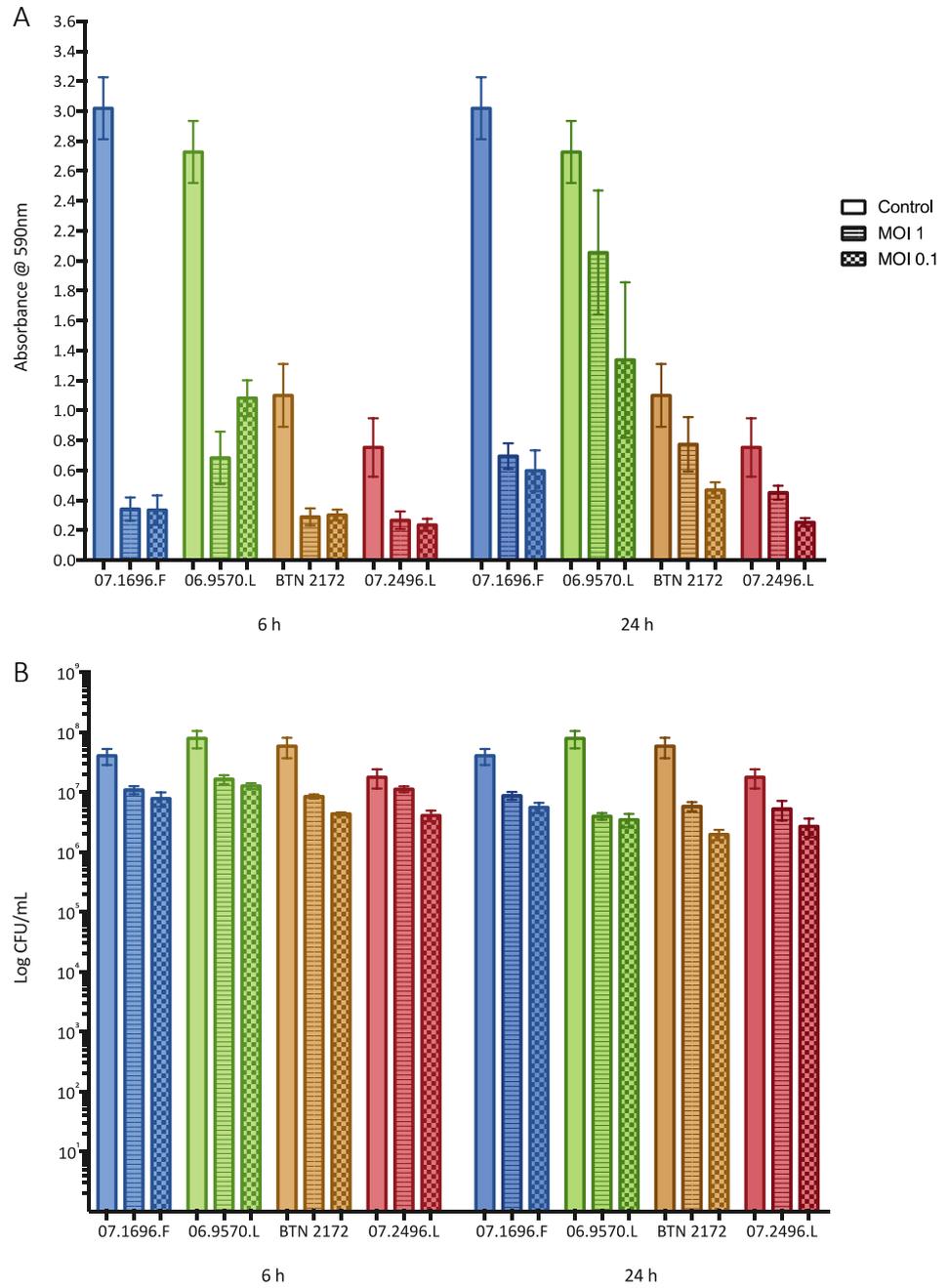
Figure 4.7: Effect of phage EW27 on mature biofilms of *S. aureus* ST36 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

4.3.6 Assessing the anti-biofilm properties of phage EW36 against ST22 and ST36 isolates

Phage EW36 was applied to mature biofilms of four ST22 (Figure 4.8) and four ST36 (Figure 4.9) isolates at two different multiplicity of infections (MOIs) of 1 and 0.1 over periods of 6 and 24 hours. Absorbance values at 590 nm following crystal violet staining of washed phage-treated biofilms revealed significant reductions in biofilm densities across all ST22 and ST36 test strains following treatment of EW36 after both 6 and 24 hours. Significant reductions in biofilm biomass was observed across all study isolates except for isolate 07.2496.L when challenged with EW36 at both MOIs ($p < 0.01$), with MOI 0.1 proving to be the most effective phage dose for degrading biofilm biomass. For both ST22 and ST36, no increase to biofilm density was observed from 6 hour to 24 hours suggesting EW36 successfully disrupted the *S. aureus* biofilms preventing biofilm regrowth. This is further supported by the greater reduction in viable cell counts when biofilms were treated for 24 hours.

Interestingly, the populations of viable bacteria recovered from each biofilm produced by the four ST22 and ST36 isolates were found to be higher in wells treated by EW36 MOI 0.1, despite producing lower absorbance readings than biofilms treated with a higher titre of phage at a MOI 1. EW36 was able to reduce viable cell numbers for both ST22 and ST36 isolates by at least one-log reduction after 6 hours and two-log reductions after 24 hours at an MOI 0.1. With biofilms treated at an MOI 1, two-log reductions were observed after 6 hours and three-log reductions after 24 hours.

135
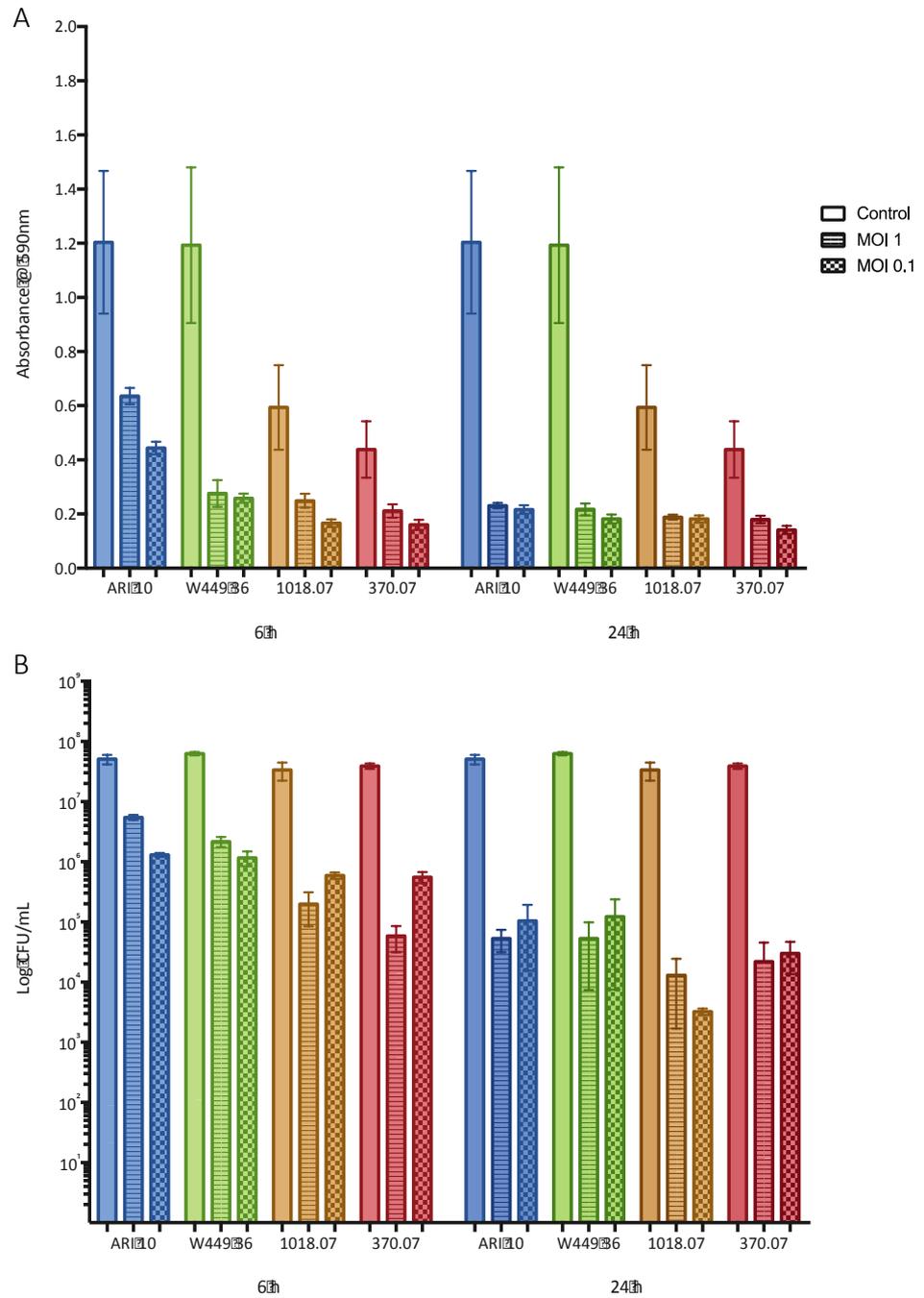
Figure 4.8: Effect of phage EW36 on mature biofilms of *S. aureus* ST22 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).
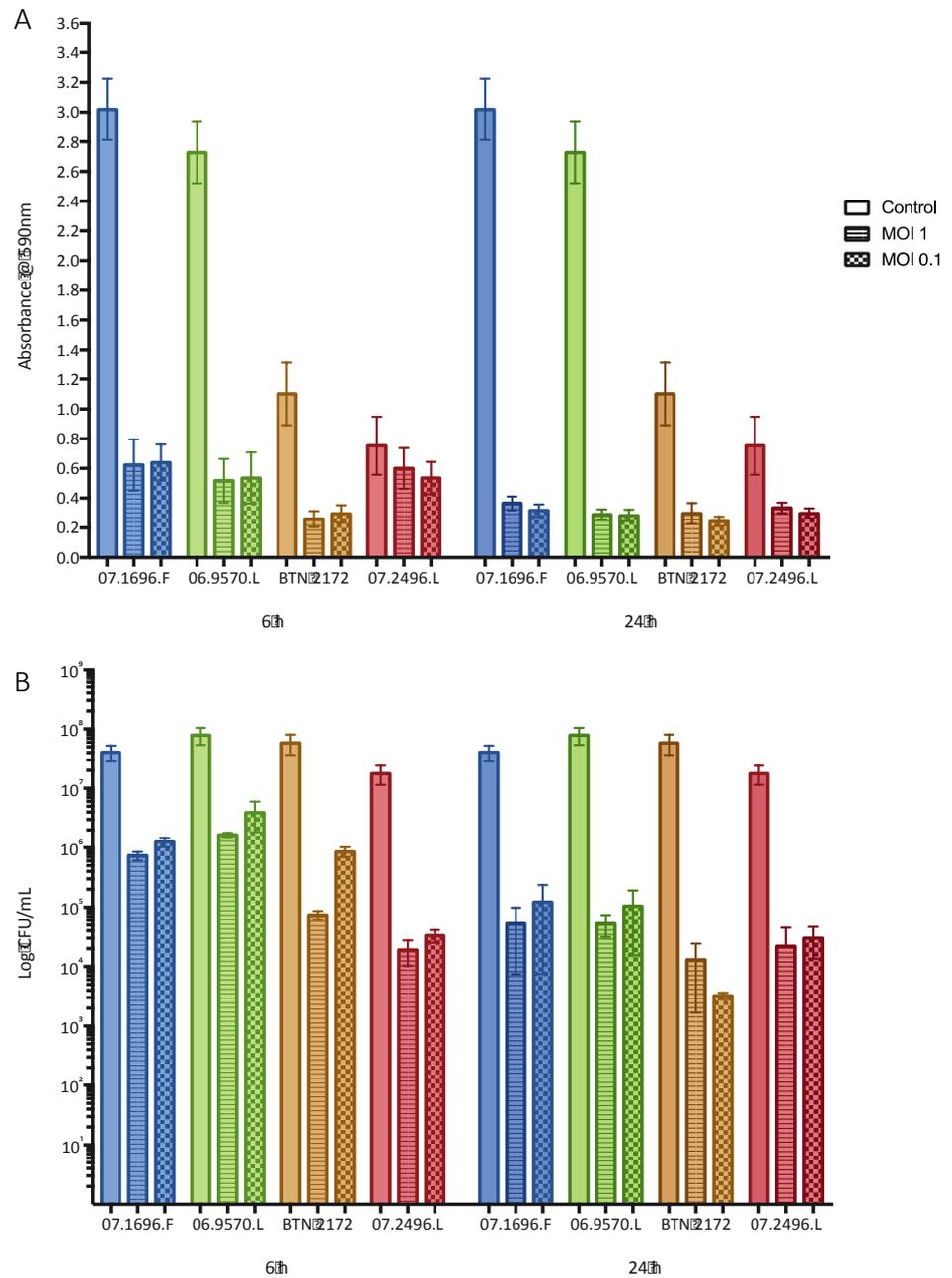
Figure 4.9: Effect of phage EW36 on mature biofilms of *S. aureus* ST36 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

4.3.7 Assessing the anti-biofilm properties of phage EW41 against ST22 and ST36 isolates

Phage EW41 was applied to mature biofilms of ST22 and ST36 isolates similarly to phage EW27 and EW36,. Significant reductions ($p < 0.01$) in biofilm biomass was observed for all ST22 and ST36 isolates tested against except for isolate 370.07 where two-log reductions in cells recovered and $60 - 93$ % reductions in biofilm biomass were observed after 6 hour treatment with phage EW41 (Figure 4.10 and Figure 4.11). Interestingly, EW41 had the least effect in reducing biofilm biomass of isolate W449 36 after 6 hours -reducing it by roughly 22 % at an MOI 1 and 19 % at an MOI 0.1, however viable cell counts were relative to all other isolates and two-log reductions were observed across both time points. Furthermore, biofilm biomass and viable cell counts recovered from the biofilms challenged with EW41 after 24 hours maintained levels similar to 6 hours exposure. Phage EW41 was able to further reduce biofilm levels of W449 36 by ~85 % when exposed for 24 hours.

ST36 biofilms challenged with EW41 for 24 hours produced higher levels of biofilm biomass and increase in cells recovered by up to one-log when compared to 6 hour exposure, suggesting regrowth had occurred within that time. Across all ST22 and ST36 isolates, both biofilm biomass and viable cells recovered were consistently lower in wells challenged with EW41 at an MOI 0.1 when compared to MOI 1, although there was no significant difference between MOI 1 and MOI 0.1.
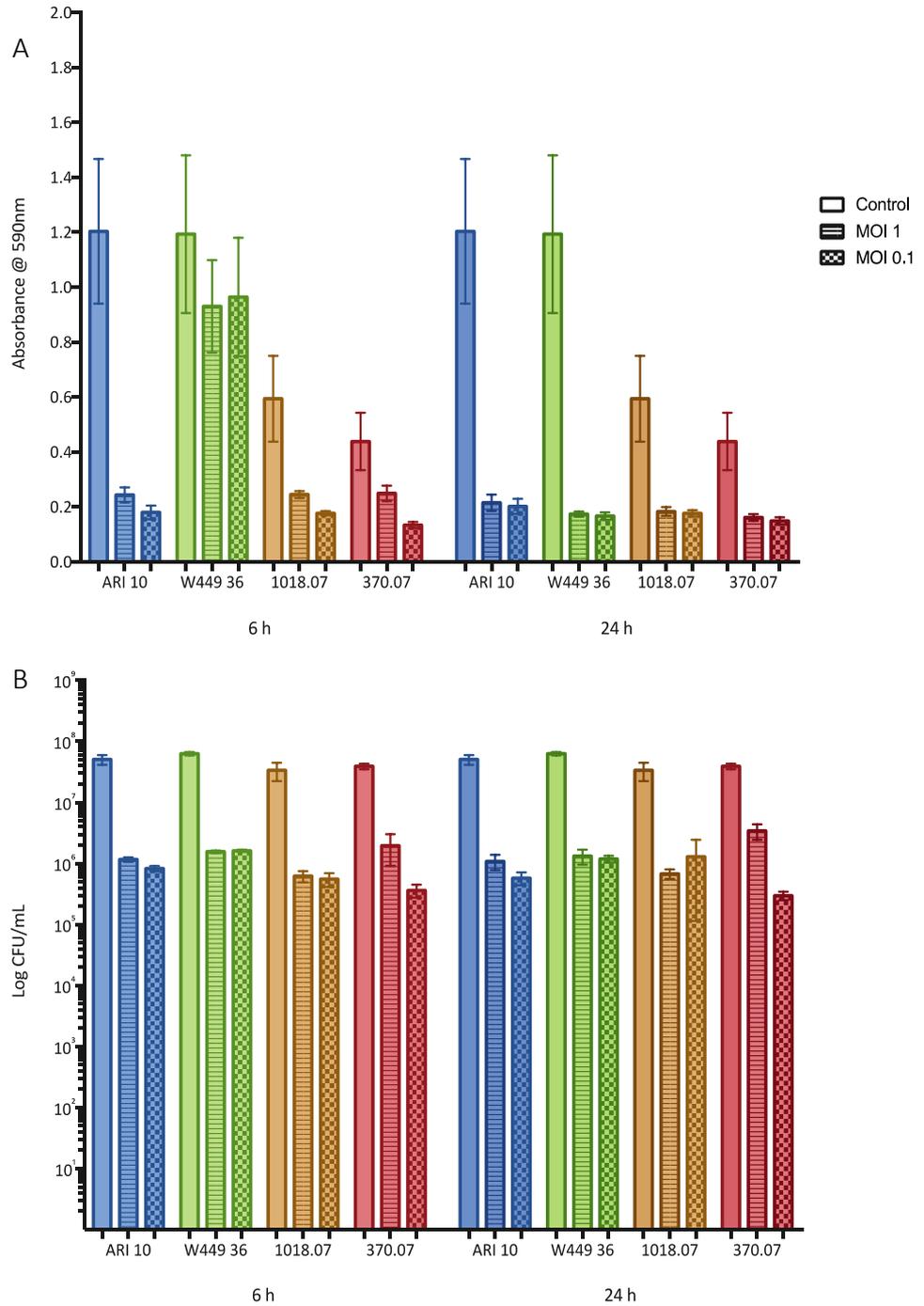
Figure 4.10: Effect of phage EW41 on mature biofilms of *S. aureus* ST22 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).
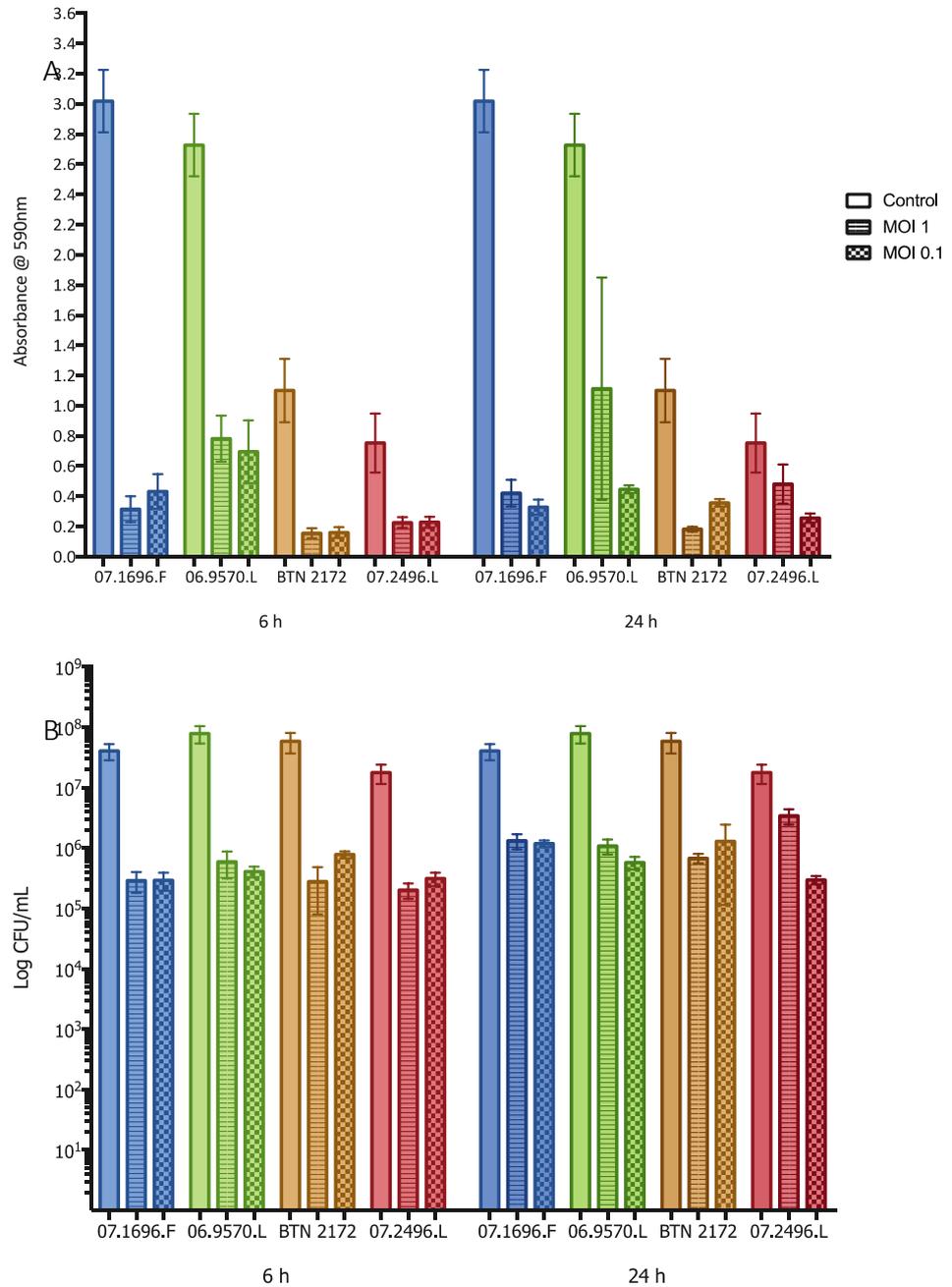
Figure 4.11: Effect of phage EW41 on mature biofilms of S. aureus ST36 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

4.3.8 Assessing the anti-biofilm properties of phage EW71 against ST22 and ST36 isolates

Phage EW71 produced the lowest values out of the four phage studied for both biofilm density and recoverable cells across all ST22 and ST36 biofilms when compared to untreated control wells (Figure 4.12 and Figure 4.13). Phage EW71 was effective at reducing ($p < 0.01$) biofilm biomass after 6 h treatment whilst greatly limiting the amount of regrowth after 24 h for all isolates tested against. Furthermore, phage EW71 was successful in reducing the number of viable cells by up to three-log after 6 h, and continued to reduce after 24 h treatment by up to four-log reductions versus those of the controls. Biofilm densities of ST22 following treatment of EW71 after 6 h ranged from 63 − 87 %, whilst consequently preventing the regrowth of all four ST22 hosts after 24 h further reducing biofilm densities.

Greater reductions in biofilm densities were also observed when ST36 isolates were challenged with phage EW71 with OD_{590} values reduced by 59 − 95 % after 6 h. Interestingly, a marginal increase in absorbance was observed in across all four ST36 biofilms when exposed to phage for 24 h. However, increases to viable cell counts were only observed for 07.1696.F and 07.2496.L suggesting phage resistance and regrowth had occurred within the two sampling periods. Phage applied to biofilms at an MOI 1 were found to be the most effective at reducing viable cell counts within the biofilm after 6 and 24 h exposures, however biofilm densities were somewhat higher with this MOI. Even so, biofilm biomasses were
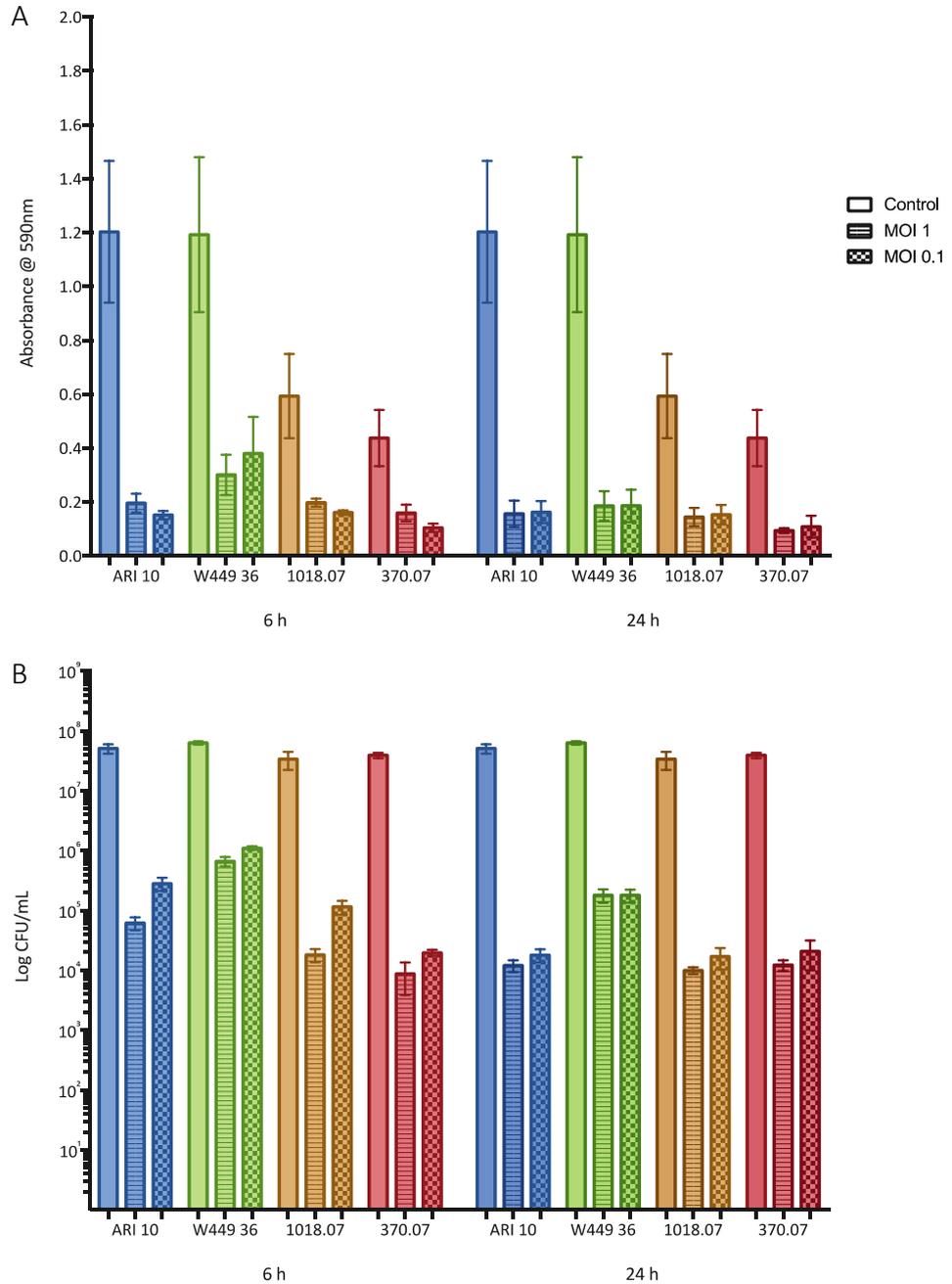
Figure 4.12: Effect of phage EW71 on mature biofilms of *S. aureus* ST22 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).
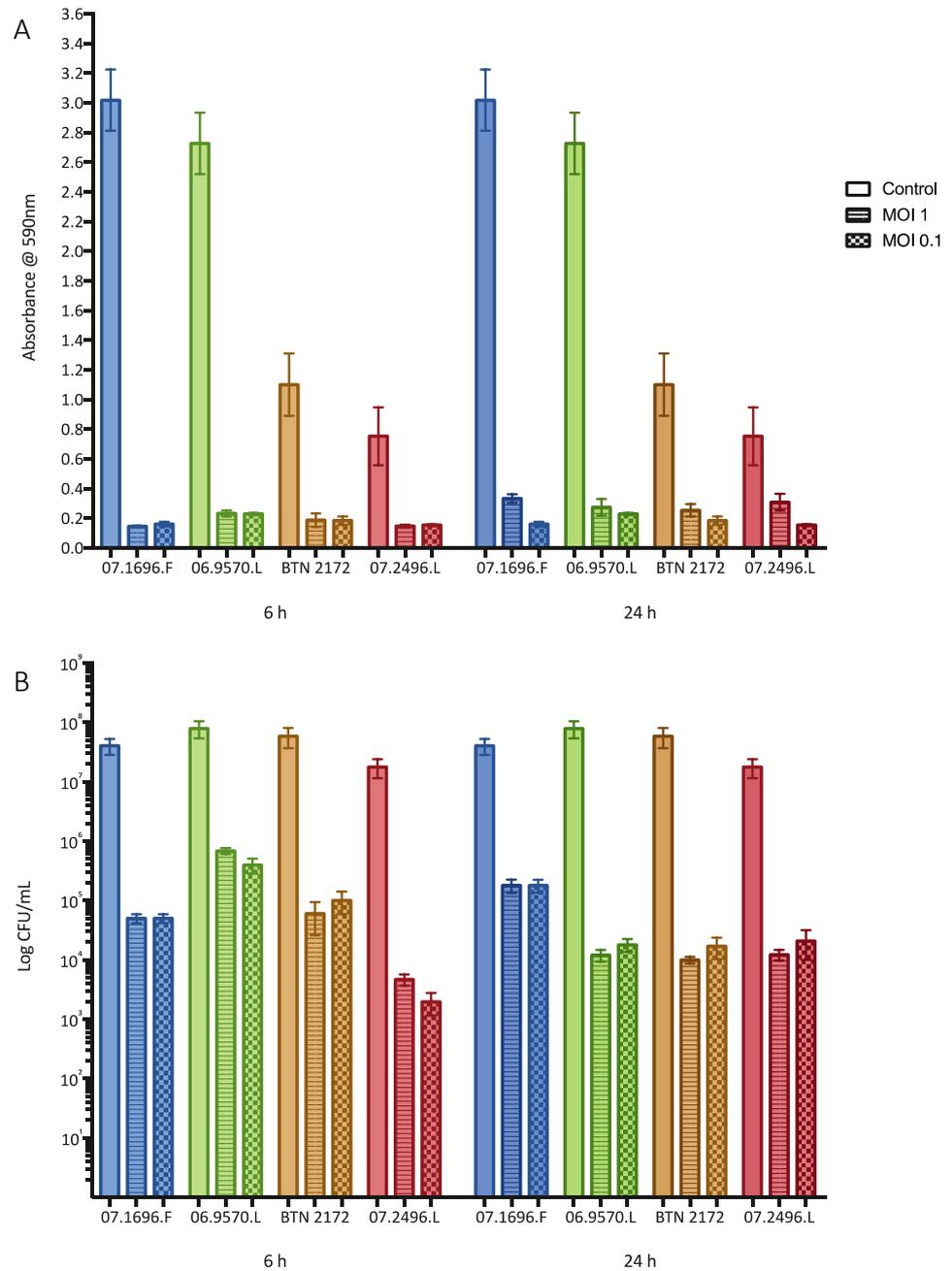
Figure 4.13: Effect of phage EW71 on mature biofilms of *S. aureus* ST36 isolates.

Static biofilms were initially grown in tissue-culture microtitre plates for 48 hours and challenged with EW36 at a multiplicity of infection of 1 and 0.1 for a period of 6 and 24 hours. A – Biofilms were initially stained with crystal violet and optical density was measured at an absorbance of 590 nm. B - Viable cells were recovered from phage treated wells by scratching and dislodging the biofilms from the surface plate wells and plated out in triplicate. Each assay was performed in triplicate, data presented as mean values (± standard deviation).

approximately similar across the majority of hosts for both 6 and 24 h treatments, except for isolate W449 36, however this difference was not significant.

4.3.9 Colony morphology

Figure 4.14 depicts heterogeneous colony morphologies that were commonly found among isolates from various phage-treated biofilms. This heterogeneity was most marked with phage EW71 treated biofilms. However, regardless of which phage the mutant had acquired resistance to, the mutants exhibited no susceptibility against the other three selected phage following spot testing on agar overlays.
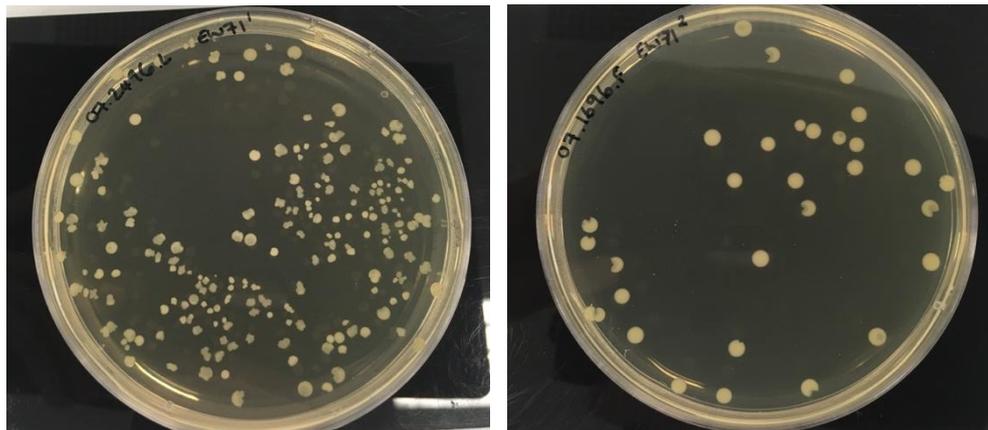
Figure 4.14: Heterogeneous colony phenotypes produced by phage-resistant mutants of 07.2496.L and 07.1696.F following exposure to EW71.

4.3.10 Confocal laser scanning microscopy of phage-treated *S. aureus* biofilms

Disruption of mature *S. aureus* biofilms grown on the surface of polystyrene tissue-culture microtitre plates was visualised by confocal laser microscopy (Figure 4.15). The quantity of live (green) and dead (red) bacteria within the biofilm was assessed using LIVE/DEAD staining. A clear reduction in biofilm biomass and living cells can

be observed after 6 h exposure to phage, with further reductions following 24 h phage treatment.
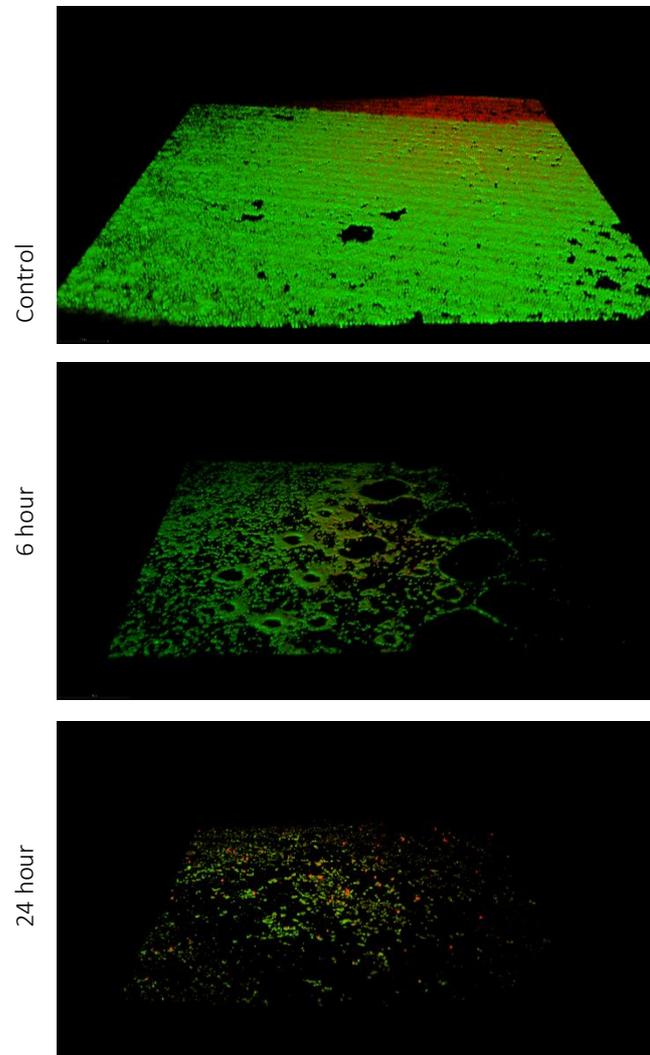
Figure 4.15: LIVE/DEAD staining of phage treated biofilms observed under confocal laser scanning microscopy (x40).

Biofilms were grown for 48 hours at 37 °C, mature biofilms where challenged with phage EW41 for 6 and 24 hour. Presence of live (green) and dead (red) bacterial cells can be observed on the surface of the polystyrene well base.

4.3.11 Evaluation of phage biofilm assays

Table 4.3 summarises the percentage reduction of *S. aureus* ST22 and ST36 isolate biofilms when challenged by EW27, EW36, EW41 and EW71 at MOIs of 1 and 0.1

after 6h and 24h. For all ST22 isolates the median reduction in biofilm biomass for MOI 1 and MOI 0.1 after 6 h was 59 % and 71 % respectively, whereas the median reduction for MOI 1 and MOI 0.1 after 24 h exposure was 72 % and 75 % respectively. Whereas, for ST36 isolates the median biofilm biomass reduction for MOI 1 and MOI 0.1 after 6 h was 80 % and 79 % respectively. After 24 h treatment, the median reduction for MOI 1 and MOI 0.1 for ST36 isolates was 75 % and 80 % respectively.

Overall, the highest biofilm biomass reductions after 6h and 24 h phage treatment was observed with phage EW41 and phage EW71 respectively, with an MOI of 0.1. Although significant reductions in biofilm biomass was observed when treated with both MOIs of phage, the greatest reductions across all ST22 and ST36 isolates after 6 and 24 h exposure was achieved when biofilms were treated at an MOI 0.1. Although each phage was able to disperse the biofilms of all study isolates,

Table 4.3: Summary table showing the relative difference in biofilm reduction of study phage at two multiplicities of infection (MOI) against four ST22 and four ST36 isolates. 48 hour biofilms were challenged with phage for 6 and 24 hours, percentages are based on control values.

	ST22	OD$_{590}$	EW27		EW36		EW41		EW71	
			MOI 1	MOI 0.1	MOI 1	MOI 0.1	MOI 1	MOI 0.1	MOI 1	MOI 0.1
6 Hour	ARI 10	1.2034	-44%	-61%	-47%	-63%	-80%	-85%	-84%	-87%
	W449 36	1.1929	-71%	-78%	-77%	-78%	-22%	-19%	-75%	-68%
	1018.07	0.5941	-59%	-74%	-58%	-72%	-59%	-70%	-67%	-73%
	370.07	0.4381	-43%	-68%	-52%	-63%	-43%	-69%	-63%	-76%
24 Hour	ARI 10	1.2034	-58%	-55%	-81%	-82%	-82%	-83%	-87%	-86%
	W449 36	1.1929	-68%	-77%	-82%	-85%	-85%	-86%	-84%	-84%
	1018.07	0.5941	-20%	-49%	-68%	-69%	-69%	-70%	-76%	-74%
	370.07	0.4381	-25%	-43%	-59%	-68%	-63%	-66%	-78%	-75%

	ST36	OD$_{590}$	EW27		EW36		EW41		EW71	
			MOI 1	MOI 0.1	MOI 1	MOI 0.1	MOI 1	MOI 0.1	MOI 1	MOI 0.1
6 Hour	07.1696.F	3.0191	-89%	-89%	-79%	-79%	-90%	-86%	-95%	-95%
	06.9570.L	2.7274	-75%	-60%	-81%	-80%	-71%	-74%	-91%	-92%
	BTN 2172	1.0997	-74%	-72%	-76%	-73%	-86%	-85%	-83%	-83%
	07.2496.L	0.7536	-64%	-69%	-20%	-29%	-70%	-70%	-81%	-79%
24 Hour	07.1696.F	3.0191	-77%	-80%	-88%	-90%	-86%	-89%	-89%	-95%
	06.9570.L	2.7274	-25%	-51%	-89%	-90%	-59%	-84%	-90%	-92%
	BTN 2172	1.0997	-30%	-57%	-73%	-78%	-83%	-68%	-77%	-83%
	07.2496.L	0.7536	-40%	-66%	-56%	-60%	-36%	-66%	-59%	-79%

complete elimination of cells was not observed across any of the hosts at either MOI as cells were recoverable when exposed to phage for 6 and 24 h treatments.

4.4 Discussion

The experiments detailed in this chapter aimed to investigate the efficacy of four lytic staphylococcal bacteriophage in disrupting and degrading *S. aureus* biofilms in an *in vitro* static biofilm system.

Biofilm-associated *S. aureus* infections are a major threat within healthcare, complicated further when caused by highly resistant MRSA isolates that make antimicrobial treatment with conventional antibiotics more difficult. The success of MRSA infections within healthcare is largely associated with (but not limited to) the emergence and spread of five globally predominant clonal lineages. Isolates from two of these successful lineages, CC22 and CC30, became the dominant MRSA within UK hospitals. ST22 and ST36 MRSA isolates (from CC22 and CC30 clonal lineages) were responsible for 95.6 % of MRSA cases between the early 1990s to mid-2000s (61, 62). Through our collection of genome-sequenced MRSA isolates, we were able to assess the biofilm forming ability of 43 ST22 and 27 ST36 isolates to identify the strongest and weakest biofilm producers in our static biofilm system.

The static biofilm model system was implemented to study the *in vitro* formation and development of *S. aureus* ST22 and ST36 isolate biofilms. This model is simple and relatively inexpensive and is the more commonly used experimental approach

for such studies, as it offers the ability for high throughput analyses and enables the rapid assessment of biofilm formation of various bacterial species, proving to be extremely useful with assays with multiple changing variables and growth conditions [223,313]. However, the closed-system approach does have its limitations, with finite supply of nutrients, these assays are prone to nutrient depletion and exhaustion [35]. Biofilm formation is directly correlated to the availability of nutrients and once depleted it can greatly limit the biofilm development. Additionally, lack of aeration and accumulation of metabolites can also be detrimental to biofilm growth [314]. To alleviate this issue, the replenishment of media is recommended and in these studies this was conducted following 24 hours incubation [315].

The formation of *S. aureus* biofilms on polystyrene surfaces relates to the ability of the bacteria to adhere to abiotic and biotic surfaces mediated by microbial surface components recognising adhesive matric molecules (MSCRAMMs) [316]. The production of polysaccharide intercellular adhesin (PIA) by *ica* operon-operated enzymes is the one of the best understood mechanisms for cell to cell adhesion and contribution to biofilm formation among staphylococci [317]. However, expression of *ica* operator across *S. aureus* clinical isolates has been found to be tightly controlled within *in vitro* conditions [46,318,319]. Despite this, a number of factors are known to influence PIA production thus leading to increased biofilm production *in vitro*, including response to anaerobic growth conditions and the presence of sugars in growth media [320–322]. The supplementation of growth medium with 1 % D-(+)-glucose greatly induced biofilm forming abilities of all our

study isolates, which is consistent with previous *S. aureus* biofilm studies [223,319,321,323].

A direct comparison of ST22 and ST36 isolates in this study revealed significant variation in the level of biofilm produced as seen in Figure 4.2 and Figure 4.3, with members of ST36 exhibiting significantly greater biofilm production. Such findings were also observed with members of CC30 (ST36) proving to be prolific biofilm producers compared to members of CC22 (ST22) under similar growth conditions, whilst demonstrating a strong correlation with the level of biofilm formation associated with specific lineages [224]. Reports on the degree of biofilm production, composition or expression levels of genes responsible for stages of biofilm formation can be associated with members of a particular ST lineage among *S. aureus* [324]. However differences in experimental conditions and the methods employed throughout various biofilm studies can make it difficult when trying to select the appropriate growth conditions for a biofilm study. In this study differing growth conditions were tested to optimise biofilm production in microtiter plates. Compared to ST36 isolates, the majority of ST22 isolates were found to be weak biofilm producers, as only 11 were considered above the nominal threshold selected to indicate good biofilm formation. However, only two ST22 isolates produced biofilm densities similar to the biofilms formed by that of the least productive ST36 isolates. So for a more accurate comparison between the two lineages, two weak biofilm producers from ST36 were selected for the phage-biofilm studies. Interestingly, although the average optical density (absorbance at 590 nm) was 1.92 across ST36, more than 6x larger than the top two ST22, the

number of viable cells recovered from each of the biofilms were considerably similar to each other between all isolates.

The four *S. aureus* phage studied in this chapter were selected based on their lytic potential demonstrated in Chapter 3. All four phage exhibited generally high efficacy in effectively reducing biofilm biomass and cell numbers of each *S. aureus* isolate after 6 and 24 h. Cell regrowth was detected following 24 h infection with phage compared to 6 h by at least one of the hosts suggesting growth of phage-resistant mutants had occurred, however this was at the expense of biofilm regrowth. These observations were similar to those reported in a previous study of *S. aureus* biofilm [190]. The observation of biofilm regrowth and phage resistance still remains a major issue and is something regularly observed in biofilms when are challenged by single lytic phage that promotes mutant selection [212]. This necessitates a phage combination approach using phage cocktails or co-administration with antibiotics to prevent the emergence of phage-resistant mutant bacteria. Previous studies have made use of the disruptive ability of phage to reduce biofilm structures produced by *S. aureus* and reduce bacterial populations enough to facilitate the penetration of antibiotics and eradicate infection [219,325]. Previous evidence suggests that the phage resistance phenotype increases sensitivity to antibiotics and also results in a loss of fitness [326]. Additionally, the application of phage cocktails consisting of multiple polyvalent phage that target different receptor proteins to prevent multi-resistance, but also increase the rate of killing, thus greatly reducing the probability of hosts acquiring resistance to phage [327]. The anti-biofilm capabilities and broad

host range demonstrated by the four study phage make them promising candidates for possible future combination studies.

Previous phage/host studies have demonstrated that by increasing the concentration of phage-to-bacteria (MOI), which, essentially increases the rate of collisions between phage and biofilm cells leads to increased rate of bacterial killing [268,328]. Additionally, greater reductions could have been facilitated by direct bacterial lysis with the possible 'lysis from without'. However there was no significant difference between MOI values, suggesting that an increased phage-to-bacteria ratio offered no advantages in reducing biofilm biomass, as described previously [328]. Overall, reductions in biofilm biomass (OD_{590}) were generally higher in biofilms treated at an MOI 0.1 (compared to 1) using any of the four phage after both 6 and 24 h treatments. The effectiveness of the low MOI demonstrates the self-perpetuating nature of lytic phage to proliferate in number, therefore only requiring small initial doses.

As mentioned above, fitness costs associated when acquiring phage-resistance, can often lead to various structural and morphological changes [329,330]. One approach is concealing the surface receptors that the phage used as docking sites to adsorb to their hosts, however these sites are often used for the uptake of nutrients [331], thus possibly limiting their growth and virulence which may be why cell numbers were able to recover after 24 h treatment compared to 6 h, yet biofilm densities remained considerably low.

Compared to the characteristic smooth, round colonies phenotypes typically produced by *S. aureus*, the recovery of heterogenous morphotypes produced by phage-resistant derivatives following phage exposure were regularly detected during this study, although most commonly observed in biofilms treated with EW71. These irregular shaped colonies have been reported in previous phage challenges and are thought to be caused by a subpopulation within that colony that has reverted back to a phage-sensitive phenotype, subsequently leading to cell death as the colony is forming [331–333]. However, the observations of 'pacman-like' colonies have not been as well documented.

Chapter 5

Genomic diversity of *Twortvirinae* infecting *S. aureus*

5.1 Introduction

Advancements in next generation sequencing (NGS) technology has greatly improved the reliability, speed and accuracy of sequencing entire genomes over recent years. The relative ease and reduction in the cost of sequencing has led to an increase in bacteriophage genomes being deposited on to databases such as National Centre for Biotechnology information (NCBI), RefSeq and the European Nucleotide Archive (ENA), and as of May 2018 there are currently >9000 complete (or near complete) sequenced phage genomes available on the NCBI website (http://millardlab.org/bioinformatics/bacteriophage-genomes/). An increasing ability to perform high throughput NGS thus further increasing the number of sequenced novel bacteriophage genomes made available on public databases will greatly improve our understanding of their complexity and evolutionary relationships with their hosts [228]. Double-stranded DNA (dsDNA) tailed phage make up the majority of sequenced bacteriophage genomes (>95 %) [334]. This increasing amount of genomic data has recently led to the re-evaluation of the bacteriophage taxonomic classification framework and the proposition of unifying classical and genomic approaches [112,335]. The order of *Caudovirales* consisting of four families of tailed bacterial viruses *Myoviridae, Siphoviridae, Podoviridae* and the new family *Herelleviridae*, greatly exceeds any other group of bacterial viruses in both phage abundance and diversity. The recent reorganization of *Caudovirales* and its subfamilies resulted in the movement of *Spounavirinae* and spouna-like viruses in to five subfamilies within the new *Herelleviridae* family: *Bastillevirinae, Brockvirinae, Jasinkavirinae, Spounavirinae and Twortvirinae* [131].

Phage genomes display extensive mosaicism and are largely modular, pointing to an evolutionary history driven by horizontal gene transfer of functional modules and non-homologous recombination [111,336] This makes it extremely challenging for species definitions and characterisation of phage population structures. Evidence suggests that phage within the same genus or type, share relatively high homology with other members of that genus, implying very limited genetic exchange between the gene pool of these phage, with differences mainly due to single nucleotide polymorphisms, insertions and/or deletions [337]. Previous comparative studies have largely focused on phage infecting pathogenic bacterial species, collections of phage from a single family, or phage belonging to all families infecting the same bacterial strain, which include phage exhibiting lytic and lysogenic lifestyles [220,283,338–343].

5.1.1 Staphylococcal phage

Many of the virulent phage infecting *S. aureus* are members of the *Twortvirinae* subfamily comprising of five genera (*Kayvirus*, *Sepunavirus*, *Silviavirus*, *Twortvirus* and unclassified *Twortvirinae*) representing as distinct clades in Figure 1.11 [112,211]. The majority of staphylococcal phage belonging to the *Twortvirinae* subfamily and display a relatively broad host range against diverse isolates from various clonal lineages, demonstrating great potential for use in phage therapy. These types of phage as well as their lytic enzymes have been studied previously [147,187,283]. However when compared at the proteome and nucleotide level, they can share little homology among themselves. Phage genome comparisons are sometimes difficult due to the limited availability of robustly annotated genes.

Many annotated phage genomes contain high levels of hypothetical or unknown proteins and in some cases, these may comprise >50 % of genes identified [344]. More accurate and thorough gene annotations will allow for better comparisons and understanding of the phage pan-genome and the distribution of shared genes within a study group. The identification and annotation of genes is crucial in order to gain a greater knowledge and understanding of the phage. Identification of these genes can lead to define the nature of the expressed protein, when and where the gene will be expressed and their genetic diversity among a large data set.

All phage lyse their bacterial hosts and in most cases, this is achieved through a holin-endolysin system [345]. These phage-encoded enzymes act to lyse their bacterial host cell causing the release of viral progeny. In *S. aureus* they degrade the peptidoglycan layer that is the main component of their cell wall. Despite their conserved biological function, endolysins are extremely architecturally diverse and vary in length and size [345]. This chapter examines the genomic differences and similarities of a collection of lytic phage infecting *S. aureus* belonging to the *Twortvirinae* subfamily and the distribution of their lytic enzymes. We previously isolated, sequenced and assembled the genomes of 22 novel bacteriophage that displayed a broad host range against a large panel of highly diverse *S. aureus* isolates. We compared these genomes to the 38 *Twortvirinae* infecting *S. aureus* genomes in the GenBank database. This allowed us to classify our 22 novel phage based on sequence homology with currently classified phage to the genus level.

We also characterized the diversity and distribution of lytic enzymes genes amongst our phage genomes and compared these to those in the public domain.

5.2 Methods

All complete *S. aureus* infecting bacteriophage genomes belonging to the *Twortvirinae* subfamily were retrieved from GenBank (https://www.ncbi.nlm.nih.gov/genome) and the European Nucleotide Archive (ENA) databases in February 2018. Each phage was individually identified as to its morphology and family by literature searching. Open reading frames (ORFs) were predicted and annotated with Prokka v1.12 [238] by searching through a user-provided protein database, in our case the Phaster viral protein database which consists of ~260,000 phage proteins (http://phaster.ca/databases) [346]. Any unannotated sequences were then searched using BLAST+ and BLASTp. However if no match was found, CDS were subsequently labelled as 'hypothetical protein'. The generated general feature files (gff) were visualised using Artemis [236]. Phage genomes were compared using DNA sequence similarity algorithm Mashtree [241] and plotted on FigTree [242].

Mash was used to compute all-against-all pairwise distances among the phage genomes, using parameters similar to those employed by homology based Average Nucleotide Identity (ANI) measures.

The output Mash matrix was visualised using heatmaply v0.14.1 [244]. Clusters were identified based on sequence similarity and visualised using BRIG v0.95 [248], individual dendrograms were created using Mashtree and pan-genomes were created for each cluster as mentioned above. Afterwards, construction of a pan-genome using the multiple sequence alignment program (MAFFT) option was achieved with Roary v3.10.2 [246]. The pan-genome was visualised with Phandango [247]. Annotated genomes were split in to individual sequence (fasta format) files and examined using HHsearch [239] (>80%, E-value < 0.01) to predict the function and structure of all identified genes in order to identify lytic enzymes homologues based on profile similarity. HH-suite utilises the hidden Markov model (HMMs) tool employed by HHpred to query translated sequences against three dimensional protein structures currently in the Protein Data Bank (PDP). This approach is uniquely different to other sequence search methods as it predicts the protein structure using homology modelling, providing greater insight in to the predicted function of each protein conserved among sequences within a cluster.

5.3 Results

In order to further characterise our 22 staphylococcal phage that we isolated and sequenced, we compared their genome sequences to the complete genomes of 38 phage infecting *S. aureus* belonging to the *Twortvirinae* subfamily downloaded from GenBank and public databases, to give a total number of 60 genomes for analysis. All 60 phage genomes were re-annotated to ensure uniformity, however, there are a number of viral protein databases available that vary in size and the

quality of their annotation such as ACLAME [347], RefSeq (ftp://ftp.ncbi.nih.gov/refseq/release/viral/viral.1.protein.faa.gz), Phaster (https://phaster.ca/databases), RAST (http://rast.theseed.org/FIG/rast.cgi), UniProt (ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/), Caudovirales (http://millardlab.org/bioinformatics/lab_server/phage-genome-annotation/).

Comparisons of these databases were performed to assess which database successfully and evenly annotated our phage genome collection. Although some databases managed to annotate a number of phage considerably well leading to higher amount of coding regions of a gene – coding sequence (CDS) with a function, some phage were poorly annotated evidenced by a high number of hypothetical protein predictions. This could be partly due to the number that included outdated annotations from older phage sequences retrieved from GenBank. Among the EW phage sequenced in this study, hypothetical proteins accounted for 61 – 67 % of the predicted genes, whereas with phage downloaded from public databases hypotheticals were less prevalent accounting for 45 – 70 % of predicted genes. A number of these original annotations were still hypothetical for example 'ORF083 staphylococcal phage G1' was the top gene shared across the phage collection. The protein was searched using BLASTp and shared 96% sequence homology (E-value 5.9e-112) to a holin gene. Furthermore, a number of protein databases were found to contain older versions of phage genomes that are poorly annotated, mainly consisting of hypothetical proteins

Table 5.1: Characteristic properties of 60 *Staphylococcus* phage genomes belonging to the *Twortvirinae* subfamily used in this study.

The 22 novel phage isolated in this study were compared with 38 complete genome sequences currently deposited on public databases, these phage were selected based on their ability to infect *Staphylococcus aureus* isolates.

Phage	Genome Size	GC Content (%)	Predicted Genes	Gene Density (genes per kb)	Average Gene Size (bp)	Percentage Coding (%)	Reference
EW1	140906	30.83	209	1.48	599	89	This study
EW2	140906	30.83	207	1.46	605	89	This study
EW3	143283	30.78	204	1.42	630	90	This study
EW4	140906	30.83	209	1.48	601	89	This study
EW5	143283	30.78	205	1.43	628	90	This study
EW6	143283	30.79	205	1.43	626	90	This study
EW7	143288	30.78	205	1.43	628	90	This study
EW9	141953	30.78	203	1.43	632	91	This study
EW13	145736	30.73	206	1.41	635	90	This study
EW15	135802	29.89	186	1.36	625	86	This study
EW18	143287	30.28	220	1.53	584	90	This study
EW20	135874	29.88	186	1.36	627	86	This study
EW22	135820	29.87	185	1.36	631	86	This study
EW26	143240	30.28	220	1.52	580	89	This study
EW27	135801	29.89	186	1.36	626	86	This study
EW29	143240	30.27	220	1.50	585	90	This study
EW36	139881	31.4	203	1.45	621	90	This study
EW41	132999	29.97	183	1.37	624	86	This study
EW42	139874	31.39	205	1.46	601	88	This study
EW71	139939	30.43	216	1.54	583	90	This study
EW72	143240	30.28	220	1.53	585	90	This study
EW74	139896	30.43	215	1.53	581	89	This study
676Z	148564	30.41	233	1.54	572	88	JX080302.2
812	142096	30.4	221	1.53	585	89	NC_029080.1
A3R	141018	30.5	212	1.48	594	88	JX080301.2
A5W	145542	30.42	232	1.57	563	88	EU418428.2
Fi200W	148481	30.39	233	1.54	571	88	JX080303.2
G1	138715	30.39	214	1.49	596	89	NC_007066.1
GH15	139806	30.23	214	1.51	596	89	NC_019448.1
K	148317	30.39	212	1.59	557	90	KF766114.1

Phage	Genome Size	GC Content (%)	Predicted Genes	Gene Density (genes per kb)	Average Gene Size (bp)	Percentage Coding (%)	Reference
ISP	138339	30.41	215	1.51	591	89	FR852584.1
JD007	141836	30.37	217	1.54	582	89	NC_019726.1
IME-SA118	139750	30.32	204	1.49	595	89	KR902361.1
IME-SA119	141028	30.33	209	1.51	586	88	KR908644.1
MCE-2014	141907	30.38	213	1.47	606	89	NC_025416.1
MSA6	131007	30.24	204	1.57	562	88	JX080304.2
P4W	147590	30.39	232	1.55	568	88	JX080305.2
P108	140807	30.22	232	1.63	543	88	NC_025426.1
phiIPLA-RODI	142348	30.42	226	1.49	597	89	NC_028765.1
phiSA012	142094	30.31	207	1.47	614	90	NC_023573.1
pSa-3	137836	30.4	209	1.49	598	89	KY581279.1
qdsa001	135563	29.86	199	1.38	609	84	KY779848.1
qdsa002	142499	30.26	231	1.56	574	89	KY779849.1
S25-3	139738	30.22	206	1.48	606	90	NC_022920.1
S25-4	132123	30.31	184	1.4	635	89	NC_022918.1
SA5	137031	30.42	198	1.57	558	88	JX875065.1
SA11	136326	30.34	186	1.37	632	87	NC_019511.1
Sb-1	127188	30.48	168	1.38	642	88	NC_023009.1
StAP1	135502	30.28	192	1.38	619	85	KX532239.1
Staph1N	145647	30.41	232	1.57	564	88	JX080300.2
Stau2	133798	29.97	178	1.33	647	86	KP881332
Team1	140903	30.33	217	1.53	585	88	NC_025417.1
Twort	130706	30.26	195	1.35	653	88	NC_007021.1
CG	142934	30.51	224	1.53	587	90	KY794641.1
Clo6	143734	30.86	213	1.47	611	90	KY794642.1
S24	139997	30.86	209	1.48	610	90	KY794643.1
LM12	143652	30.25	227	1.53	584	89	MG721208.1
Remus	134643	29.97	174	1.35	639	86	NC_022090.1
Romulus	131332	30.01	164	1.33	653	87	NC_020877.1
fRuSau02	148464	30.22	236	1.55	564	87	MF398190.1

such as staphylococcal phage G1, leading to large set of genes were simply annotated as 'ORF G1'. To prevent this, specific annotations such as the hypothetical phage G1 proteins were removed from the database and all phage were subject to re-annotation, resulting in an increase in number of genes annotated with a predicted function. Each phage was annotated using Prokka against the Phaster protein database, the genomes of these phage ranged from 127,188 bp to 148,564 bp. The average G+C % content of the *S. aureus* host is 32.9 %, whilst *S. carnosus* possesses the highest of all sequenced staphylococcal genomes with 34.6 %. The average G+C % content of the phage is 30.5 ± 1.1 %, a lower percentage than their bacterial hosts.

5.3.1 Phylogenetics

The collection of 60 phage genome sequences were compared through pairwise sequence comparisons using Mash to study their phylogenetic relationship, the resulting dendrogram was plotted using Figtree and exported as a Newick tree. The dendrogram resolved several distinct clusters that are highlighted below (Figure 5.1), these clusters represent closely-similar phage sharing a high degree of homology with those of other study genomes. Long branch lengths between clusters show how divergent some clusters are from others, whilst close relationships between phage members within each cluster can be observed. The clusters containing the phage sequenced in this study were further analysed. It also highlights a number of singletons that exhibit very little sequence homology to the other phage with no close genetic relatives, including the notable phage Twort.
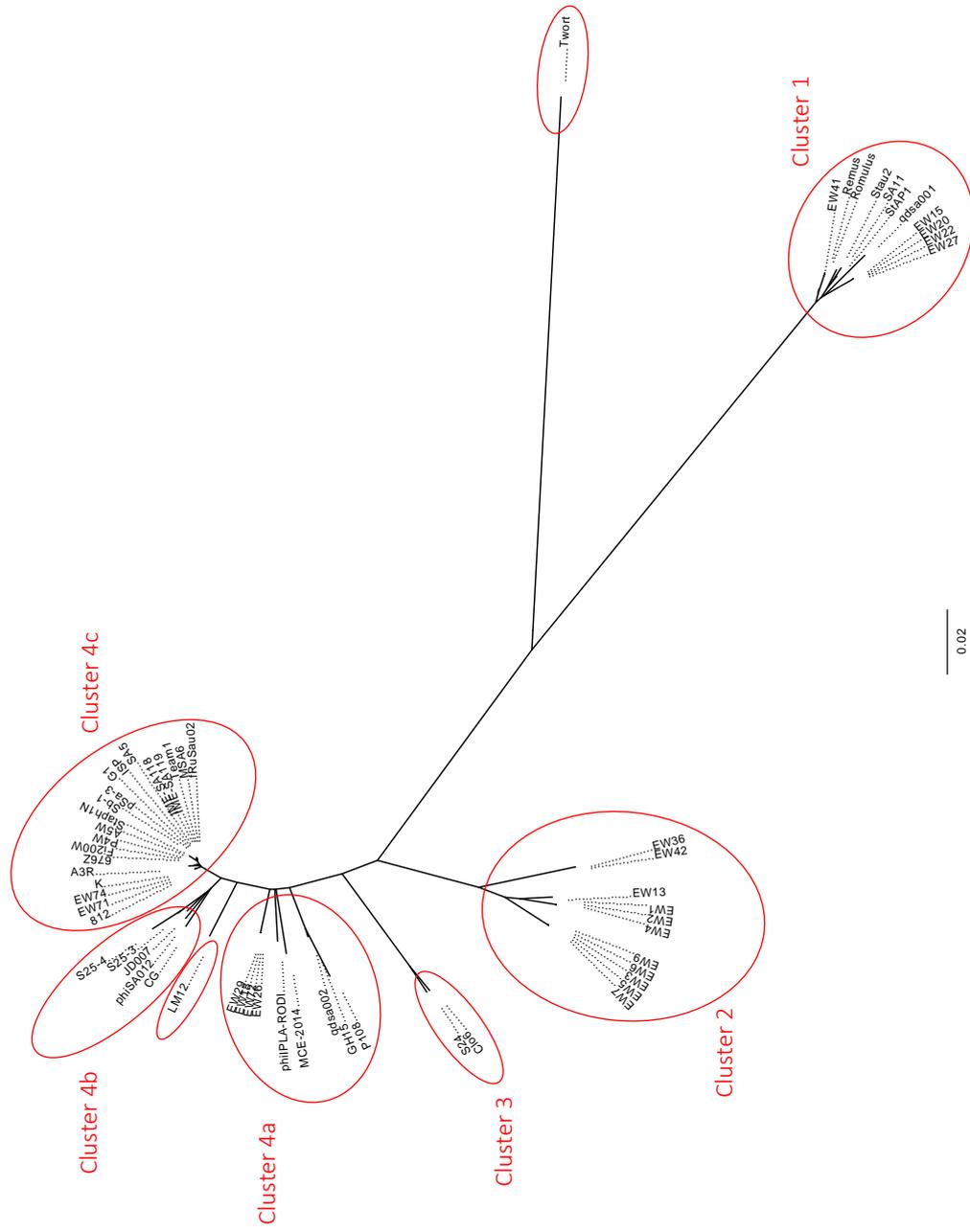
Figure 5.1: Dendrogram illustrating the nucleotide sequence similarities of 60 *Staphylococcus* bacteriophage genomes belonging to the *Twortvirinae* subfamily.

This collection compares our 22 sequenced S. aureus phage genomes and 38 publicly available genome sequences from GenBank and the European Nucleotide Archive. Clusters 1-4 represent phage sharing high sequence similarities. These groupings are largely arbitrary but made on the basis of MiniHash distance estimates.

5.3.2 Pan-genome analysis

A total of 1634 ORFs of at least 122 bp were predicted within all 60 phage genomes.

Of these, 380 ORFs (23 %) had predicted gene functions based on the Phaster

database, analysis of phage sequence annotations revealed the percentage DNA that

codes for proteins (percentage coding) ranged from 84 to 91 % with phage qdsa_001

possessing the lowest and phage EW9 the highest, with an average number of genes

per kilobase of 1.41.

In order to characterise the core and accessory genome amongst all 60 phage, the pan-

genome pipeline tool Roary was used. Interrogation of the pan-genome identified no

predicted genes that were found to be present in all phage genomes, with the highest

number of phage that were found to share the same gene was 48. In total, 31 genes

were shared across 48 phage and further analysis identified that these genes consisted

of mainly, structural proteins associated with DNA repair, replication and tail

structures. A number of unidentified CDS genes (hypothetical proteins) that were

shared across phage present within distinguishable groups were individually examined

using UniProt BLASTp (http://www.uniprot.org/blast/) in order to assign their function.

This helped to identify a number of genes that had originally been overlooked during

initial annotation with Prokka. The proportion (%) of accessory genes that are not

found present across all of the phage genomes can be designated according to Roary

as i) 'soft-core' - genes found in ≥95 % to <99 % of genomes, ii) 'shell' genes : ≥15 % to

<95 %) or iii) 'cloud' genes: present in <15 % of genomes. When all 60 phage are

compared there are a considerable number of accessory genes (<95 %) that are unique

and specific to a single phage and these are for the most part annotated as hypothetical

163

or unknown function (Table 5.2). Furthermore, a large number of unidentified ORF's (initially annotated as hypothetical proteins) that were found to be shell genes, were individually examined using UniProt BLAST option in order to assign their function. Gene by gene comparisons to identify homologues sharing high synteny allowed us to identify a large number of hypothetical proteins previously overlooked during initial annotation. This was mainly achieved using comprehensive sequence and structural homology searches using BLASTp and HHpred tools respectively. From this, we were able to identify genes based on their predicted protein structures that were present across each phage. The pan-genome was visualised with Phandango, whose input is a Roary 'gene presence/absence' tab separated values text file and a Newick format dendrogram file, in this case a Mashtree dendrogram, to examine the distribution of proteins in relation to DNA sequence similarity is show in Figure 5.2.

Table 5.2: Number and distribution of genes in the core and accessory genome across members of each cluster.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4a	Cluster 4b and 4c
Core	120 (47%)	109 (29%)	163 (65%)	140 (40%)	113 (25%)
Soft-core	0	0	0	20 (5%)	0
Shell	99 (38%)	229 (61%)	87 (35%)	19 (28%)	178 (40%)
Cloud	39 (15%)	35 (10%)	0	94 (27%)	154 (35%)
N = genes	258	373	250	351	445
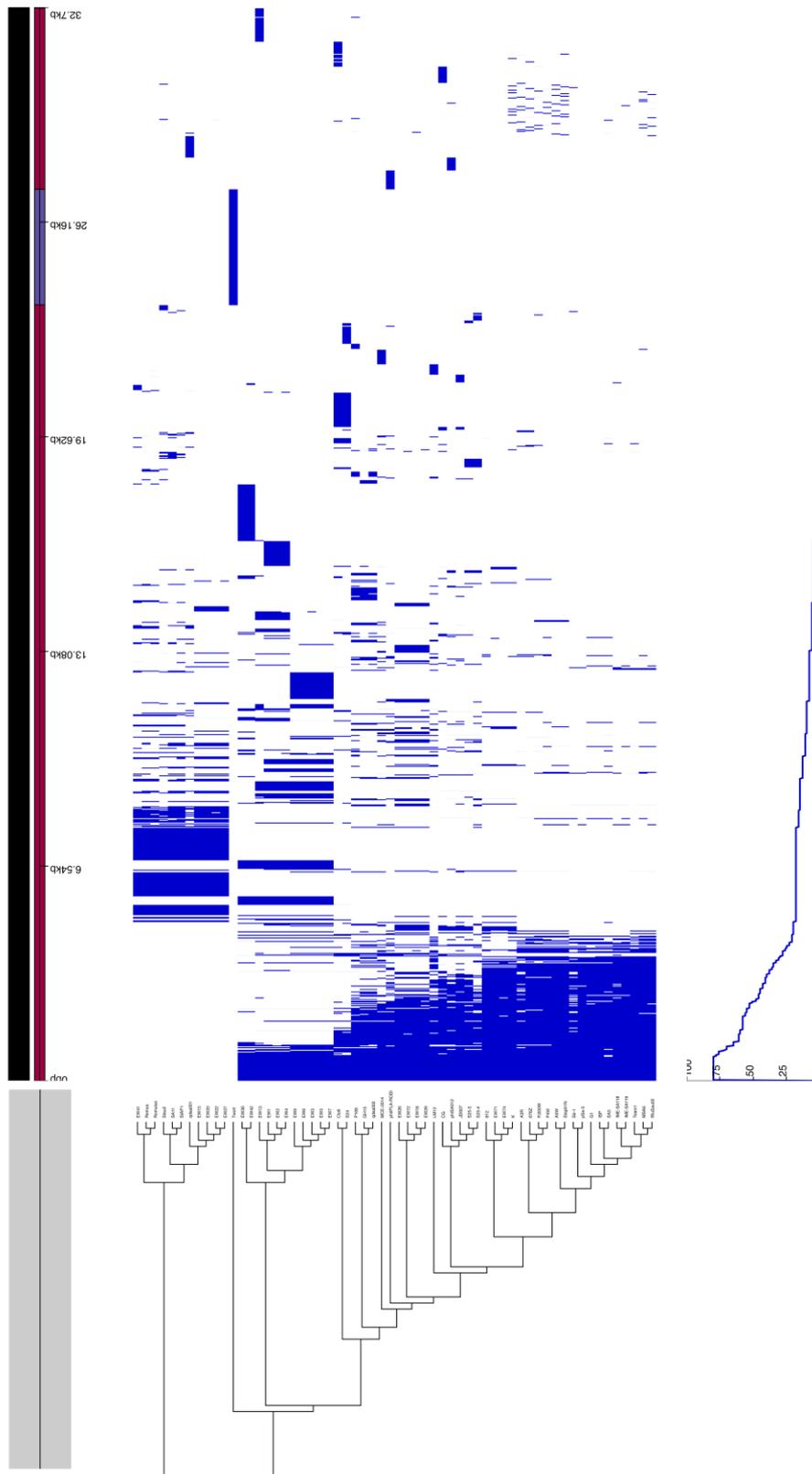
Figure 5.2: Pan-genome of 60 staphylococcal phage genomes belonging to the *Twortvirinae* subfamily.

Blue blocks represent genes that are present in the phage genome in order to identify genes conserved within clusters of phage and understand their diversity.

Dendrogram was generated using Mash distances, Roary was used to visualise gene presence/absence across the phage collection.

5.3.3 Comparative genomics

Pairwise sequence similarities of all phage genomes plotted against each other were computed using Mash distances. The Mash matrix was read in Rstudio and plotted as a heat map using heatmaply (Figure 5.3). Mash pairwise values were also converted in to percentages, values among phage sharing a high percentage identity were coloured green and red if phage were more divergent, allowing the distinction of phage species. Of the 60 sequenced phage, 58 were organised in to 4 clusters, with one cluster consisting of only two phage members, and Cluster 4 being sub-divided into a further 3 sub-clusters.

These phage were grouped into clusters based on genome synteny using an all-vs-all approach, high similarity (1 = 100 %) is indicated by yellow and blue indicated low similarity (0.7 = >70 %). Several clusters and two singletons are clearly visible and these are labelled in Figure 5.3 and correspond to the clusters identified in Figure 5.1. These distinct clusters observed were found to reflect current ICTV genera, 1 – *Silviavirus*, 2 - Unclassified EW, 3 – Unclassified *Kayvirus* subgroup, 4 – *Kayvirus* and two singleton clusters Twort the lone member of *Twortvirus* and LM12 which appears to share a higher degree of similarity to *Kayvirus* members.

Interestingly, our sequenced phage (EW1 – EW22) were present in all three major clusters, of which, one in particular Cluster 2 was found to possess exclusively phage from this study. A number of the phage from this study displayed extensive similarity up to 100 % similarity with phage K and other EW phage. Members of Cluster 3 shared the highest similarity of 91 – 92 % to those in Cluster 4, whose two members Clo6 and

S24 are currently placed in an unclassified phylogenetic subgroup of *Kayvirus* sharing 92 % sequence similarity with key members of the *Kayvirus* genus found in Cluster 4. Cluster 2 shares 89 – 90 % homology with phage of Cluster 4, and considerably different to *Silviavirus* Cluster 1 with only 75 – 76 % nucleotide similarity. Of which are highly divergent sharing only 75 – 77 % with all other clusters, however Twort phage still remains the only member of its genus and its incredibly divergent genome shows no levels of synteny with other staphylococcal phage.

The distinguishable clusters visible on the heatmap (Figure 5.3) are congruent to those identified in both the dendrogram (Figure 5.1) and pangenome (Figure 5.2). It was clear that members of cluster 1 and phage Twort exhibit relatively low homology with members of the remaining clusters, and may suggest why there are no core genes found shared across all phage genomes. During analysis, we found that members within clusters that were previously classified correlated to the phage genus previously assigned. Clusters were further interrogated in order to understand the pan-genomes and similarities of study phage and members within those groups. Not only did this also help assign unclassified phage to fall in to their respective genus, it helped determine where our 22 study phage could belong and also understand more about them in relation to already classified phage.
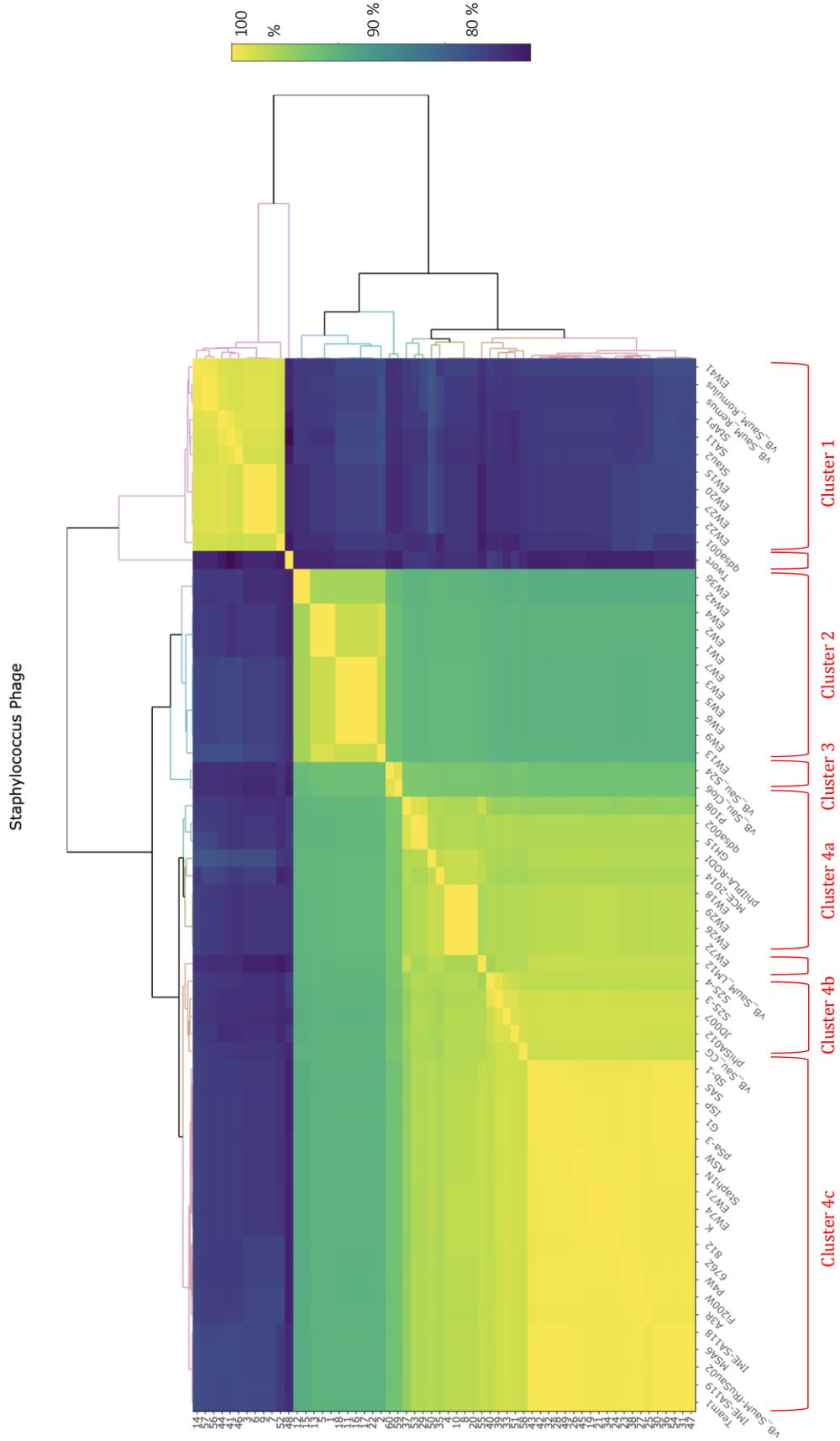
Figure 5.3: All-vs-all pairwise comparisons of 60 staphylococcal phage genomes.

Matrix was generated using Mash and visualised with Heatmaply, the degree of genome synteny is indicated by the colour gradient from high (yellow) to low (blue).

5.3.4 Cluster 1

Cluster 1 contains 11 phage, of which, five members were from this study. Pairwise comparisons for each phage using an all-against-all approach identified the five EW phage to have a high degree of coverage (>95 %) with members of *Silviavirus* between 97.46 % and 100 % (Table 5.3), sharing significantly lower homology to members of *Kayvirus* (~75 %) and *Twortvirus* (~73 %). EW phage were compared against the genomes of current members within the cluster stAP1, qdsa001, SA11, Stau2, Romulus and Remus. When compared to the Remus, the type strain for *Silviavirus* genus, the EW phage demonstrated a high degree of synteny - EW15 (98.5 %), EW20 (98.5 %), EW22 (98.5 %), EW27 (98.5 %) and EW41 (99.3 %) and should be rightfully considered as new additions to this genus. However, we note that a number of EW phage share a significantly high sequence identity to each other – of up to 100 %, indicating that they are essentially the same phage that were most likely present in numbers across multiple ASP2 tanks at the Davyhulme wastewater treatment site as a number of

Table 5.3: Pairwise comparisons of phage infecting *S. aureus* belonging to the *Twortvirinae* subfamily.
Phage of Cluster 1 are compared against each other to determine their sequence homology, the level of identity is given as a percentage and coloured from green to red based on the degree of homology.

	SA11	StAP1	Stau2	qdsa001	Remus	Romulus	EW15	EW20	EW22	EW27	EW41
SA11	100%										
StAP1	99.27%	100%									
Stau2	99.05%	98.65%	100%								
qdsa001	97.57%	97.96%	97.46%	100%							
Remus	98.32%	98.78%	98.14%	97.87%	100%						
Romulus	98.25%	98.70%	98.11%	97.89%	99.90%	100%					
EW15	98.31%	98.70%	97.99%	97.67%	98.51%	98.47%	100%				
EW20	98.31%	98.70%	97.99%	97.67%	98.51%	98.47%	100%	100%			
EW22	98.31%	98.70%	97.99%	97.67%	98.51%	98.47%	100%	100%	100%		
EW27	98.31%	98.70%	97.99%	97.67%	98.51%	98.47%	100%	100%	100%	100%	
EW41	98.23%	98.53%	98.11%	97.94%	99.33%	99.33%	98.35%	98.35%	98.35%	98.35%	100%

samples were collected on each day.

Pan-genome analysis was performed against the 11 phage as depicted in Figure 5.4. A total of 258 predicted genes were identified across the 11 phage within Cluster 1. The group shares a core genome consisting of 120 (47 %) predicted and a shell consisting of 99 (38 %) genes and a cloud of 39 genes (15 %). Of which, 96 genes had predicted function with 162 hypothetical proteins. The core genome consisted of 120 genes with 49 having a predicted function that include virion structural proteins such as major capsid and tail proteins, lytic enzymes and transcriptional regulatory proteins.

A number of unique genes often present only in a single phage were classed as individual groups, further interrogation of these genes using BLAST and HHpred identified each of them to produce the same protein of similar structure with high identity and coverage. For example, group_38 (Roary output) only present in phage qdsa001 (qdsa001_00144) putatively encodes for a DNA polymerase I protein. To confirm this, the sequence was searched on HHpred, the top hit matched it up to a DNA polymerase (PDB ID - 1KFS_A - 99.2 %). The PDB ID 1KFS_A was searched across the HHpred outputs (hhr files) of all phage belonging to Cluster 1. Interestingly, the 1KFS_A hit was found to be the top hit for group_37 which was shared among the other 10 phage. Group_37 was initially annotated as a hypothetical protein, and is found present in all 11 phage including qdsa001, but as (qdsa001_00143). The amino acid sequence for qdsa001_00144 (group_38) was aligned against a number of protein sequences that were classed as group_37 from different phage in the cluster (Figure 5.5) using UniProt align function.
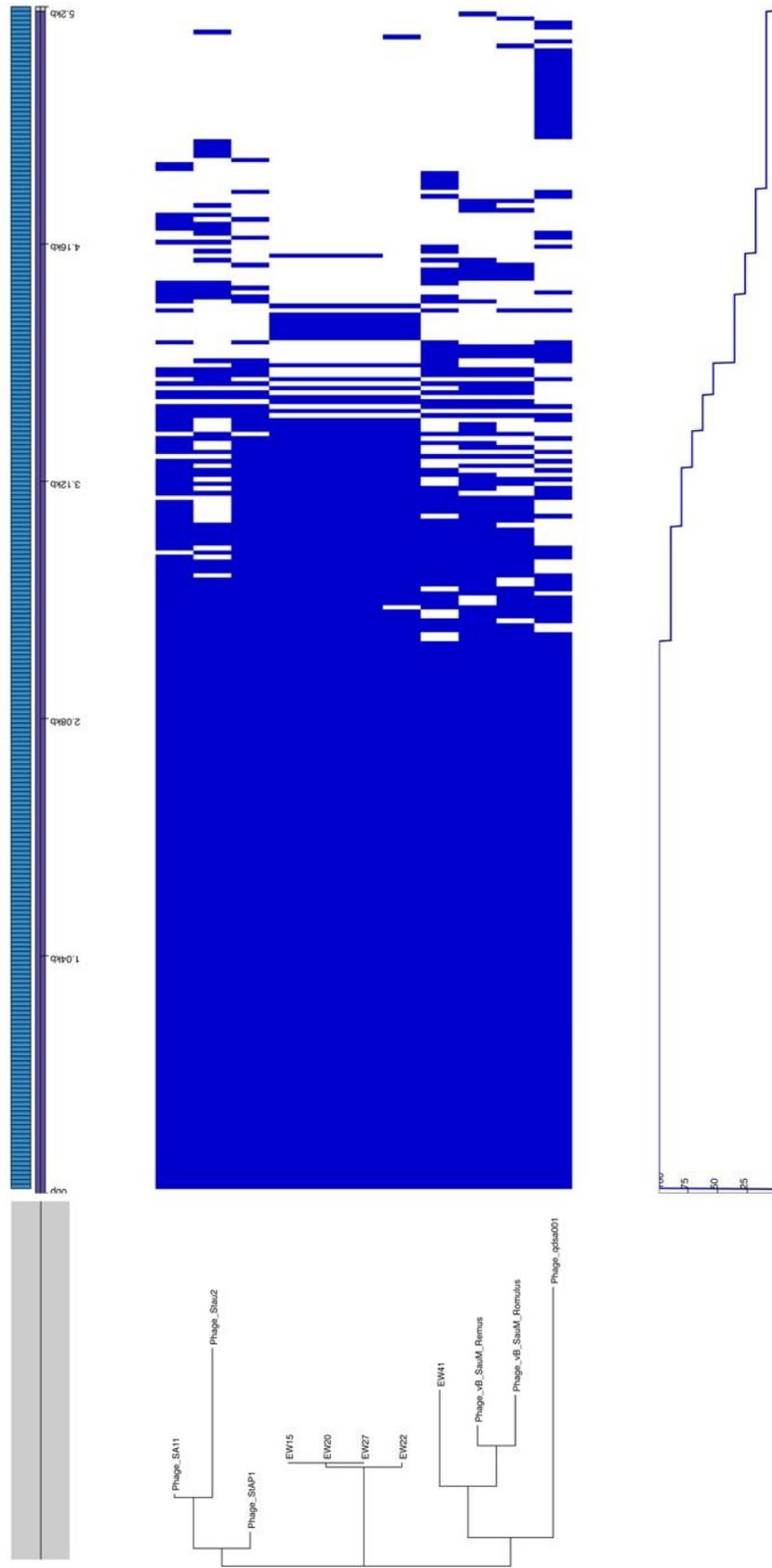
Figure 5.4: Pan-genome of all 11 genomes of phage infecting *S. aureus* belonging to the *Silviavirus* genus.

Blue blocks represent genes that are present to identify conserved and unique genes across all phage. Dendrogram was generated using Mash distances, Roary was used to obtain a gene presence/absence matrix.

Alignment identified a high similarity between these genes, except for the sequence of group_37 which contained a 1200 bp insertion (Figure 5.5). This intragenic region was picked out and manually searched through HHpred and found a 100 % hit to an endonuclease found in *Pyrococcus furiosus* (1DQ3_A). When 1DQ3_A was searched through the hhr files, it was found present in all 10 phage except for qdsa001. The genes encoding for this protein all belonged to group_207 rightly titled intron-encoded putative endonuclease, which was found shared across all 10 phage except for qdsa001. It appears that group_38 (qdsa001_00144) which is only present in qdsa001, is essentially group_37 with group_207 inserted in the middle of the gene. Inteins that interrupt protein-encoding genes are a regular occurrence in phage genomes including *Spounavirinae*, especially members of the *Silviavirus* genus [131,348], and may present a challenge with gene prediction tools suggesting possible issues with annotation, thus leading to errors within pan-genome analysis.

Figure 5.5: The location of group I intron.

The inserted endonuclease that phage qdsa001 harbours can be seen when the gene (Group_38) is aligned against the sequences of a number of members possessing Group 37 gene, and is a recognised group I intron that often encodes for endonucleases.

5.3.5 Cluster 2

The 11 members of Cluster 2 consist of only phage from this study and demonstrate a high level of shared homology (Table 5.4), with pairwise sequence alignments ranging from 95.41 % to 100 %. Examination of the Mash matrix indicates that EW1, EW2 and EW4 to be the same phage, in addition to EW3, EW5 and EW7. This can also be applied to EW36 with EW42 both being 100 % identical phage too, possibly suggesting that these phage were present in a number of ASP sewage tanks when samples were collected. It is also worth noting that, EW9 shares an extremely high similarity of 99.98 % to EW3, EW5, EW6 and EW7. The sequence similarity of EW36 and EW42 against the other phage is 95.41 %. When compared to other clusters, EW phage in Cluster 2 share a relatively high identity with members in all subgroups of Cluster 4 and Cluster 3 ranging from 89.40 % to 90.83 %, whilst sharing around 75.30 % to members of Cluster 1.

Table 5.4: Pairwise comparisons of phage infecting *S. aureus* belonging to the *Twortvirinae* subfamily.
Phage of Cluster 2 are compared against each other to determine their sequence homology, the level of identity is given as a percentage and coloured from green to red based on the degree of homology.

	EW1	EW13	EW2	EW3	EW36	EW4	EW42	EW5	EW6	EW7	EW9
EW1	100%										
EW13	98.21%	100%									
EW2	100%	98.21%	100%								
EW3	97.37%	97.43%	97.37%	100%							
EW36	95.41%	95.54%	95.41%	95.41%	100%						
EW4	100%	98.21%	100%	97.37%	95.41%	100%					
EW42	95.41%	95.54%	95.41%	95.41%	100%	95.41%	100%				
EW5	97.37%	97.43%	97.37%	100%	95.41%	97.37%	95.41%	100%			
EW6	97.37%	97.43%	97.37%	99.99%	95.41%	97.37%	95.41%	100%	100%		
EW7	97.37%	97.43%	97.37%	100%	95.41%	97.37%	95.41%	100%	100%	100%	
EW9	97.38%	97.43%	97.38%	99.98%	95.42%	97.38%	95.42%	99.98%	99.98%	99.98%	100%

Phylogenetic analysis placed these newly isolated phage in a distinct clade closest to members of *Kayvirus* genus. When searched using BLASTn these same members were among the top hits for the EW phage. The 11 EW phage of Cluster 2 were compared against the *Kayvirus* type strain phage K used as a reference using the circular genome analysis tool, BRIG (Figure 5.6). BRIG comparison illustrates the high level of synteny between EW phage shared with phage K, however also revealed a number of segments that were not present among the cluster 2 members compared to phage K.
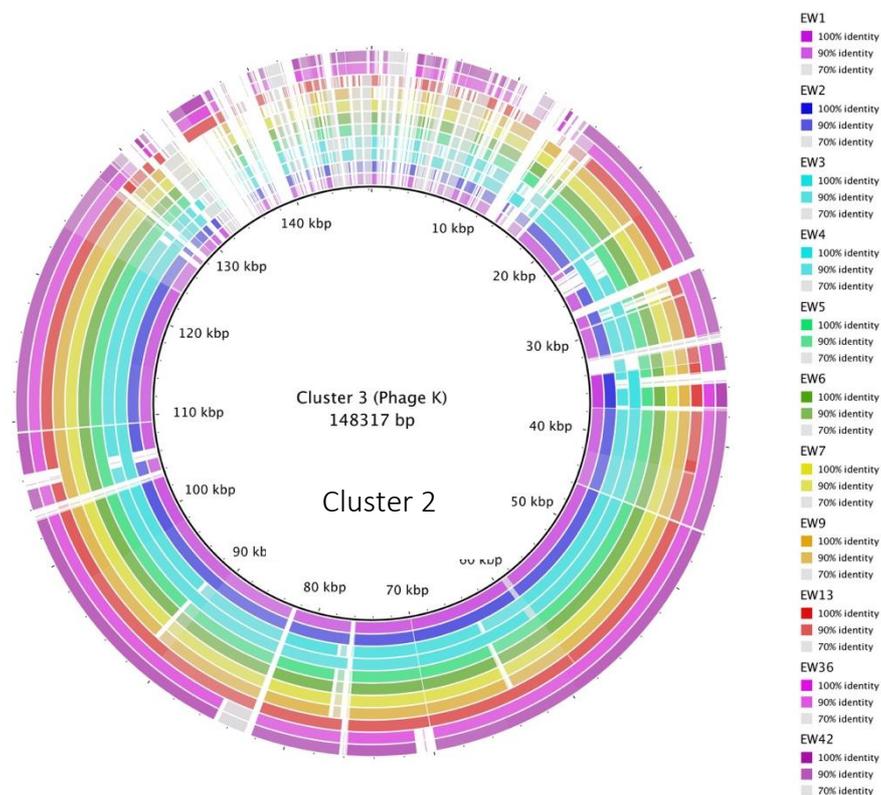
Figure 5.6: Comparison of 11 unclassified EW staphylococcal phage of cluster 2.

The BRIG image illustrates the variability among members of the *Kayvirus* genus, each coloured ring represents an EW phage genome similarity against the *Kayvirus* type strain Phage K used as a reference. The brightness and shading of the colour for each phage represents the percentage (%) similarity against the phage K genome. Percentage identity of less than 40 % are indicated by no colouration.
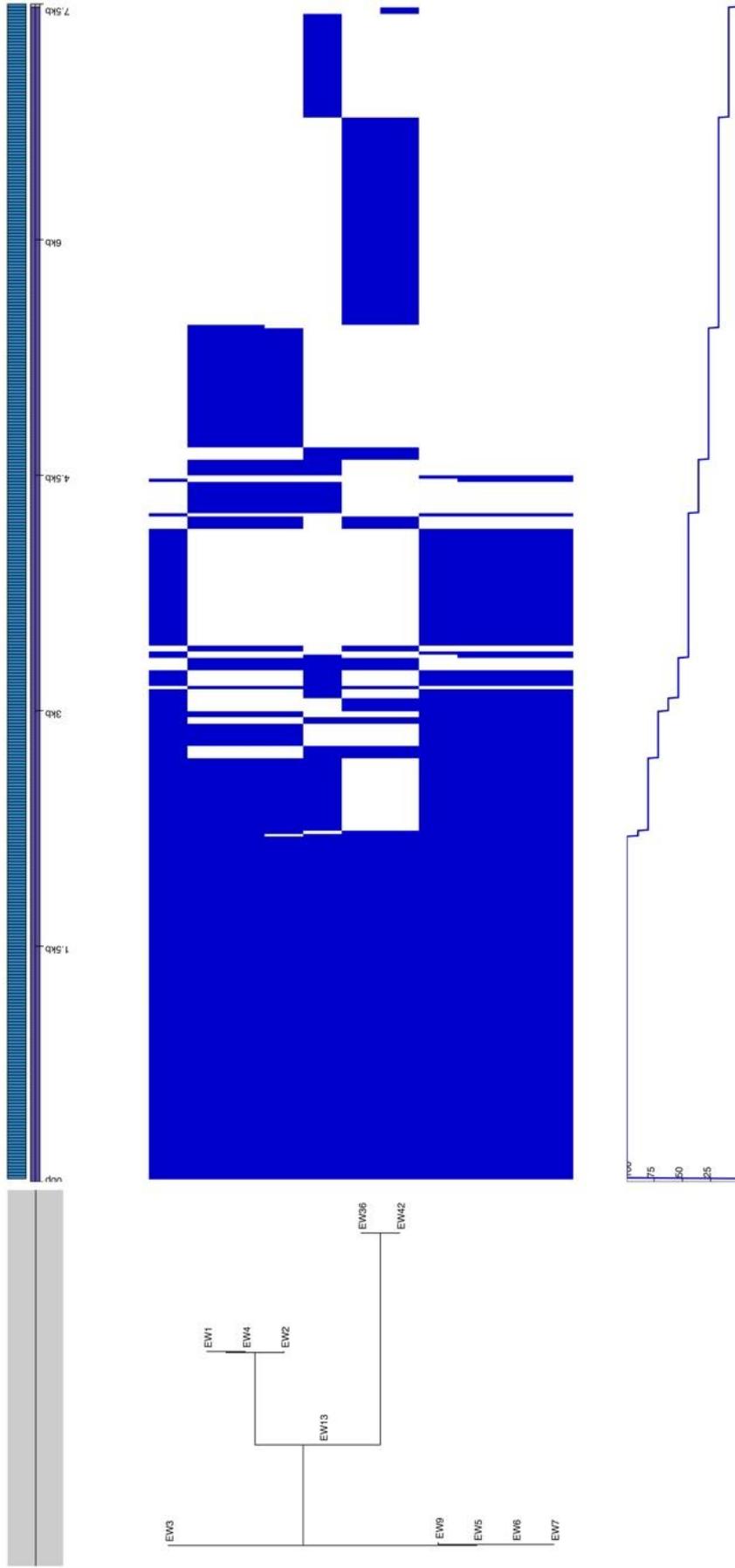
Figure 5.7: Pan-genome of 11 unclassified EW genomes of phage infecting *S. aureus*.

Blue blocks represent genes that are present to identify conserved and unique genes across all phage. Dendrogram was generated using Mash distances, Roary was used to obtain a gene presence/absence matrix.

176

Pan-genome analysis identified a total of 373 predicted genes across the 11 phage of Cluster 2 as depicted in Figure 5.7. Of which, 99 had a predicted function. The group shared a rather small core genome consisting of 109 (29 %) a shell of 229 (61 %) and a cloud genome of 35 (10 %). The hypothetical proteins present across all cluster 2 members were further interrogated using HHpred, and found that the predicted protein structures share high homology to structural virion proteins, transcriptional and interestingly several lytic enzymes. Gene-by-gene comparisons were performed in order to detect overlooked genes and determine whether genes found only in a single phage were in fact unique or due to poor annotation.

5.3.6 Cluster 4

Cluster 4 consisting of three sub-clusters 4a, 4b and 4c is made up of 35 phage genomes, 6 of which were from this study. Members belong to the biggest genus of the *Twortvirinae* subfamily known as *Kayvirus*. This cluster can be further sub-divided into 3 small clusters based on >3 % sequence similarity (Figure 5.1). However, cluster pairwise sequence comparisons revealed that phage EW18, EW26, EW29, and EW72 were 100 % identical to each other. In addition, EW71 and EW74 were found to also share 100 % similarity. Nucleotide similarities (Table 5.5) against the 29 members of *Kayvirus* identified 6 EW phage, both EW71 and EW74 to be 100 % identical to genus type strain phage K and 99.93 % to phage 812. On the other hand, phage genomes sharing the lowest similarity to EW phage were P108 (94.39 %) to EW71 and EW74. Interestingly, P108 is also the most similar genomically to phage Clo6 and S24 of Cluster 3 which have not yet been classified. When clusters are taken into consideration, members of Cluster 4a share a 1 % difference between phage in Cluster 4b, whereas

Table 5.5: Pairwise comparisons of phage infecting *S. aureus* belonging to the *Kayvirus* genus.

Phage of Cluster 5 are compared against each other to determine their sequence homology, the level of identity is given as a percentage and coloured from green to red based on the degree of homology.

	6762	812	A3R	A5W	F12D0W	G1	GH15	IME-SA118	IME-SA119	ISP	JD007	K	MCE-2014	MSA6	P108	PAW	S25-3	S25-4	SA5	Sb-1	StaphN	Team1	pSa-3	phiIPLA-RODI	phiSA012	qdsau002	fRuSau02	LM112	CG	EW18	EW26	EW29	EW71	EW72	EW74
6762	100%																																		
812.000	99.69%	100%																																	
A3R	99.68%	99.52%	100%																																
A5W	99.77%	99.69%	99.62%	100%																															
F12D0W	99.85%	99.66%	99.66%	99.74%	100%																														
G1	99.70%	99.61%	99.54%	99.77%	99.77%	100%																													
GH15	96.06%	95.96%	96.13%	96.10%	96.06%	96.14%	100%																												
IME-SA118	99.63%	99.53%	99.45%	99.69%	99.63%	99.89%	96.07%	100%																											
IME-SA119	99.63%	99.53%	99.47%	99.70%	99.63%	99.90%	96.06%	99.94%	100%																										
ISP	99.70%	99.61%	99.53%	99.77%	99.71%	99.96%	96.13%	99.89%	99.96%	100%																									
JD007	97.95%	98.01%	97.96%	98.00%	97.97%	97.97%	96.43%	97.95%	97.97%	97.97%	100%																								
K	99.63%	99.93%	99.46%	99.73%	99.59%	99.59%	96.02%	99.58%	99.59%	99.67%	98.06%	100%																							
MCE-2014	95.99%	95.87%	96.03%	96.05%	95.99%	95.99%	95.95%	95.94%	95.94%	96.01%	95.60%	95.92%	100%																						
MSA6	99.63%	99.54%	99.47%	99.69%	99.63%	99.89%	96.05%	99.91%	99.93%	99.88%	97.90%	99.58%	95.92%	100%																					
P108	94.46%	94.34%	94.48%	94.52%	94.46%	94.53%	98.33%	94.46%	94.53%	94.53%	94.57%	94.39%	96.39%	94.43%	100%																				
PAW	99.87%	99.70%	99.70%	99.80%	99.89%	99.73%	96.07%	99.66%	99.73%	99.73%	97.95%	99.64%	96.02%	99.66%	94.50%	100%																			
S25-3	97.71%	97.74%	97.76%	97.75%	97.73%	97.77%	95.92%	97.73%	97.73%	97.79%	98.61%	97.80%	95.55%	97.72%	94.50%	97.70%	100%																		
S25-4	96.87%	96.87%	97.06%	96.95%	96.89%	97.02%	96.25%	96.97%	96.97%	97.04%	97.65%	96.93%	95.80%	96.95%	95.13%	96.91%	98.95%	100%																	
SA5	99.66%	99.56%	99.49%	99.72%	99.66%	99.91%	96.04%	99.85%	99.86%	99.94%	97.98%	99.62%	95.97%	99.85%	94.49%	99.69%	97.65%	97.70%	100%																
Sb-1	99.44%	99.35%	99.41%	99.51%	99.45%	99.72%	96.12%	99.65%	99.67%	99.71%	97.86%	99.40%	95.96%	99.66%	94.44%	99.47%	97.70%	97.03%	99.47%	100%															
StaphN	99.78%	99.70%	99.62%	99.99%	99.75%	99.77%	96.11%	99.70%	99.70%	99.70%	98.01%	99.74%	96.05%	99.70%	94.54%	99.81%	97.76%	96.96%	99.81%	99.73%	100%														
Team1	99.64%	99.55%	99.47%	99.70%	99.64%	99.91%	96.08%	99.93%	99.95%	99.90%	97.97%	99.60%	95.96%	99.94%	94.48%	99.66%	97.73%	96.97%	99.66%	99.66%	99.71%	100%													
pSa-3	99.69%	99.60%	99.53%	99.76%	99.69%	99.96%	96.13%	99.88%	99.89%	99.95%	98.02%	99.65%	96.01%	99.88%	94.52%	99.72%	97.79%	97.02%	99.90%	99.74%	99.77%	99.72%	100%												
phiIPLA-RODI	96.41%	96.39%	96.46%	96.46%	96.41%	96.44%	96.50%	96.36%	96.37%	96.44%	95.98%	96.44%	96.93%	96.35%	96.26%	96.37%	95.77%	96.00%	96.40%	96.43%	96.41%	96.41%	96.42%	100%											
phiSA012	97.78%	97.86%	97.72%	97.83%	97.81%	97.80%	96.06%	97.73%	97.72%	97.80%	98.23%	97.91%	95.70%	97.71%	94.97%	97.79%	98.35%	97.35%	97.75%	97.61%	97.84%	97.75%	97.78%	95.78%	100%										
qdsau002	96.04%	95.95%	96.13%	96.13%	96.06%	96.17%	99.78%	96.08%	96.07%	96.15%	96.44%	96.01%	95.88%	96.07%	98.44%	96.08%	95.99%	96.32%	96.07%	96.15%	96.14%	96.08%	96.15%	96.46%	96.16%	100%									
fRuSau02	99.62%	99.52%	99.45%	99.69%	99.62%	99.89%	96.04%	99.92%	99.95%	99.90%	97.95%	99.58%	95.93%	99.94%	94.43%	99.65%	97.71%	96.95%	99.85%	99.65%	99.69%	99.94%	99.87%	96.34%	97.71%	96.05%	100%								
LM112	97.22%	97.15%	97.18%	97.29%	97.26%	97.29%	95.82%	97.20%	97.20%	97.29%	96.28%	97.21%	96.75%	97.17%	97.29%	97.25%	97.25%	95.32%	97.25%	97.24%	97.30%	97.30%	97.29%	96.55%	95.53%	95.72%	97.18%	100%							
CG	97.64%	97.73%	97.55%	97.66%	97.65%	97.63%	95.95%	97.57%	97.58%	97.63%	97.97%	97.75%	95.40%	97.57%	94.41%	97.64%	97.77%	96.87%	97.57%	97.44%	97.67%	97.64%	97.64%	95.77%	97.86%	95.92%	97.77%	96.18%	100%						
EW18	96.73%	96.97%	96.48%	96.75%	96.71%	96.77%	96.22%	96.22%	96.72%	96.78%	96.40%	97.02%	97.19%	96.71%	95.97%	96.73%	96.16%	96.25%	96.71%	96.44%	96.77%	96.73%	96.78%	97.46%	96.35%	96.17%	96.35%	95.73%	96.38%	100%					
EW26	96.73%	96.97%	96.48%	96.75%	96.71%	96.77%	96.22%	96.72%	96.72%	96.78%	96.40%	97.02%	97.20%	96.71%	95.97%	96.73%	96.16%	96.25%	96.73%	96.44%	96.77%	96.73%	96.78%	97.46%	96.35%	96.17%	96.35%	95.73%	95.73%	100%	100%				
EW29	96.73%	96.97%	96.48%	96.75%	96.71%	96.77%	96.22%	96.72%	96.72%	96.78%	96.40%	97.02%	97.19%	96.71%	95.97%	96.73%	96.16%	96.25%	96.73%	96.44%	96.77%	96.73%	96.78%	97.46%	96.35%	96.17%	96.35%	95.73%	96.38%	100%	100%	100%			
EW71	99.63%	99.93%	99.46%	99.73%	99.59%	99.59%	96.02%	99.58%	99.59%	99.67%	98.06%	100%	95.92%	99.58%	94.39%	99.64%	97.80%	96.93%	99.64%	99.40%	99.74%	99.60%	99.66%	96.44%	97.91%	96.01%	99.64%	97.21%	97.75%	99.60%	99.75%	100%	100%		
EW72	96.73%	96.97%	96.48%	96.75%	96.71%	96.77%	96.22%	96.72%	96.72%	96.78%	96.40%	97.02%	97.19%	96.71%	95.97%	96.73%	96.25%	96.25%	96.72%	96.44%	96.77%	96.72%	96.73%	97.46%	96.35%	96.18%	96.16%	95.70%	96.38%	96.39%	96.39%	96.39%	97.02%	100%	
EW74	99.63%	99.93%	99.46%	99.73%	99.59%	99.66%	96.02%	99.66%	99.59%	99.67%	98.06%	100%	95.92%	99.58%	94.39%	99.64%	97.82%	96.93%	99.62%	99.40%	99.74%	99.64%	99.66%	96.44%	97.91%	96.01%	99.58%	97.21%	97.21%	96.73%	96.39%	97.02%	100%	97.02%	100%

phage members of 4b share a 4 % difference between 4c, although the majority of phage share >95 % similarity amongst each other.

In total, 503 genes were predicted across the Cluster 4. Of which, 85 genes had a predicted function following annotation, although a large majority of these genes were annotated as 'ORF [Phage]' which may have impeded proper annotation. Cluster 4 phage comprise a core genome consisting of 99 (20 %) genes conserved across all phage, with only 12 with predicted function. In addition, an extensive accessory genome was present compromising of a soft core consisting of 24 genes (5 %), a shell of 144 (28 %) and a cloud of 236 (47 %) that makes up the accessory genome as demonstrated in Figure 5.8.
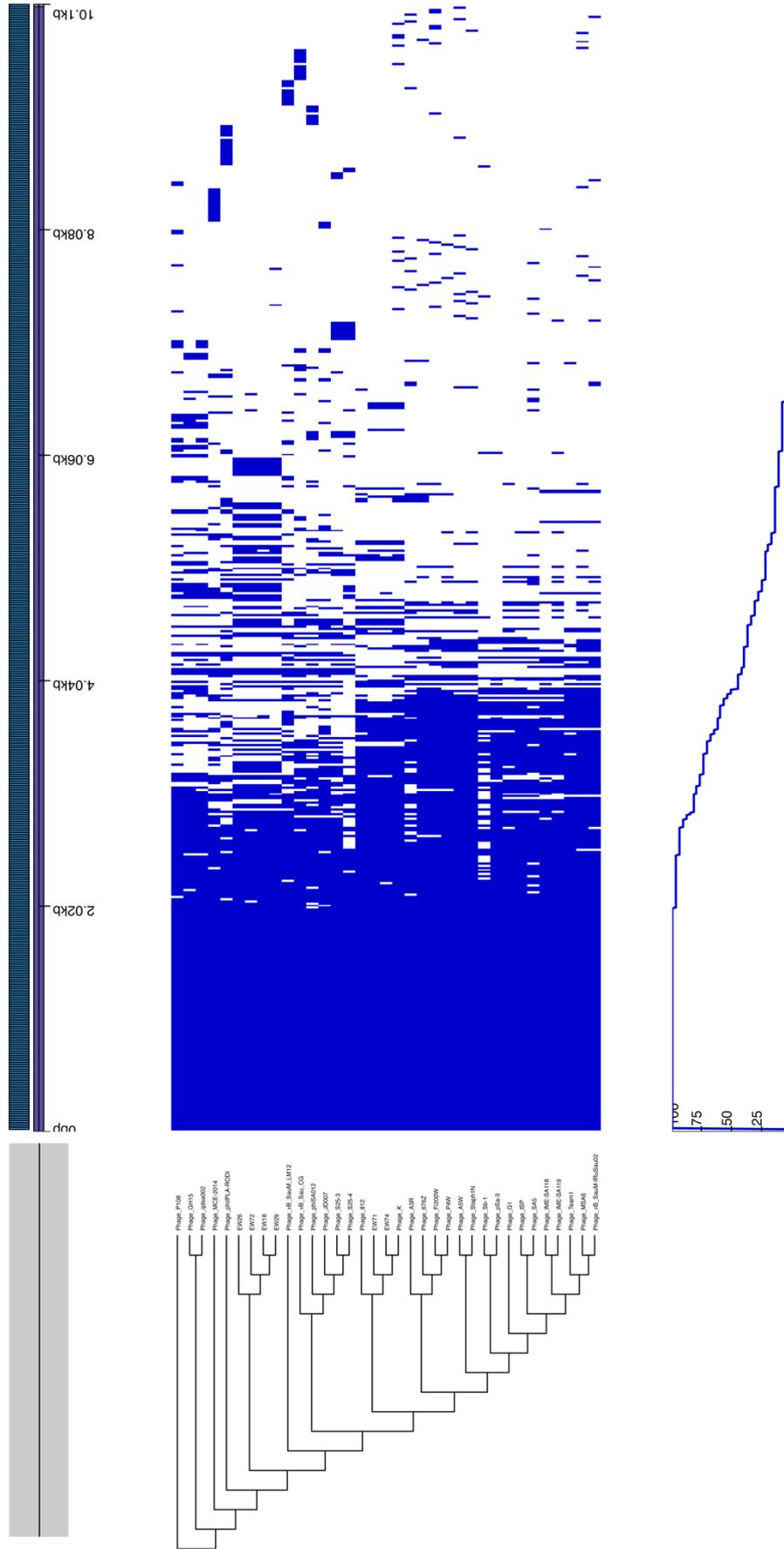
Figure 5.8: Pan-genome of 35 phage infecting *S. aureus* genomes belonging to the *Kayvirus* genus.

Blue blocks represent genes that are present to identify conserved and unique genes across all phage. Dendrogram was generated using Mash distances, Roary was used to obtain a gene presence/absence matrix

5.3.7 Lytic proteins

Individual deduced phage protein sequence were used to interrogate that of known sequences using HHpred database. This uses structural rather than sequence homology (BLAST) to search for similar proteins. We identified 407 potential genes out of 12906 by searching for structural lytic protein and enzyme homologues. We screened all matches for key words including hydrolase, endolysin, transglycosylase, N-acetylemuramoyl-L-alanine, amidase, lysozyme, autolysin, peptidoglycan, LysK, CHAP domain, endopeptidase and holin in HHpred output files to identify possible genes encoding for lytic proteins.

The 407 potential genes were individually screened to filter out low probability hits with a cut-off value of 90 % probability and an E - value of 0.0001, resulting in a final collection of 237 genes encoding for putative lytic enzymes. Once filtered, each HHpred output was aligned to their hosts and compared to the original gene presence/absence table obtained from the pan-genome analysis of all phage (Figure 5.2) in order to determine the original nucleotide sequence that each group was assigned to. The 237 genes discovered in the database fell in to 21 individual groups that were originally assigned by Roary, the sequence for each group was translated to its amino acid and aligned using MUltiple Sequence Comparison by Log-Expectation (MUSCLE). The dendrogram produced using MUSCLE was aligned against the manually curated lytic enzyme gene presence/absence using all 60 phage, that were ordered based on the Mash distance tree above. This helped detect lytic enzymes that were specific to a single cluster or shared across numerous clusters (Figure 5.9).

All 60 phage in this study possessed a gene that shares structural homology to a cysteine, histidine-dependant amidohydrolase/peptidase domain (CHAP domain) protein (2K3A_A – *Staphylococcus saprophyticus* subsp.), an endopeptidase that is found within a distinct clade in tree A, highlighted in red (Figure 5.9) among all the identified lytic enzyme genes. The CHAP domain protein that Twort possesses shares the lowest sequence similarity to members of Cluster 4. Members of Cluster 2, 3 and 4 are found to share the same CHAP domain protein (2K3A_A, group_19), further analysis of the three groups 18, 19 and 20 found that two large insertions at the start of group_18 and 19. Yet displayed 80 % similarity among all three when these insertions weren't considered. Additionally, all phage possessed a gene that shared similarity in protein structure to LysK (4CSH_A – staphylococcal phage K), a well-studied phage endolysin that has been reported to lyse live staphylococcal cultures [146]. Although no single gene was found present among all the phage used in this study, all phage carried genes encoding for lytic enzymes that shared structural similarity to the same predicted protein.

The majority of phage in each cluster share the same number of genes encoding for lytic enzymes ranging from 4 to 6. Cluster 4 displays the highest diversity amongst genes with a number of phage sharing uniquely different genes encoding for the same protein. Furthermore, a number of phage were found to possess up to 2 additional genes encoding a hypothetical protein sharing high structural similarity to an endolysin (group_10 and group_298) within their genomes. This was also found in Cluster 2 where EW1, EW2 and EW4 carrying an extra gene to other members of that cluster, which was also predicted to be an endolysin (lss).

Among the group of lytic enzymes found, HHpred identified a number of genes that shared protein structural homology to resuscitation promoting factor (4EMN_B, 4CGE_A and 2N5Z_A). It has been discovered that these resuscitation promoting factor (Rpf) proteins share a high structural similarity to lysozymes and soluble transglycosylases [349], and can be seen with group_1366, group_1636, group_329 and group_23, sharing high structural homology to group_329 and group_23 encoding for lysozymes. These Rpf-related proteins are capable of hydrolyzing the peptidoglycan motifs of the membrane [350], degradation of the peptidoglycan is required for the resuscitation of dormant cells [149]. It was also found that a specific motif within a tape measure protein of a *Mycobacterium tuberculosis* phage, sharing high structural similarity to an Rpf protein, facilitated the efficiency and infectivity of phage against stationary cells. All clusters were found to encode genes predicted to be Rpf-related proteins and lysozymes, however further analysis of the HHpred outputs among Cluster 1 identified Group_413 that shared structural similarity to a peptidoglycan hydrolase (4Q4G_X), also shares high similarity to a resuscitation promoting factor protein (3NE0_A).
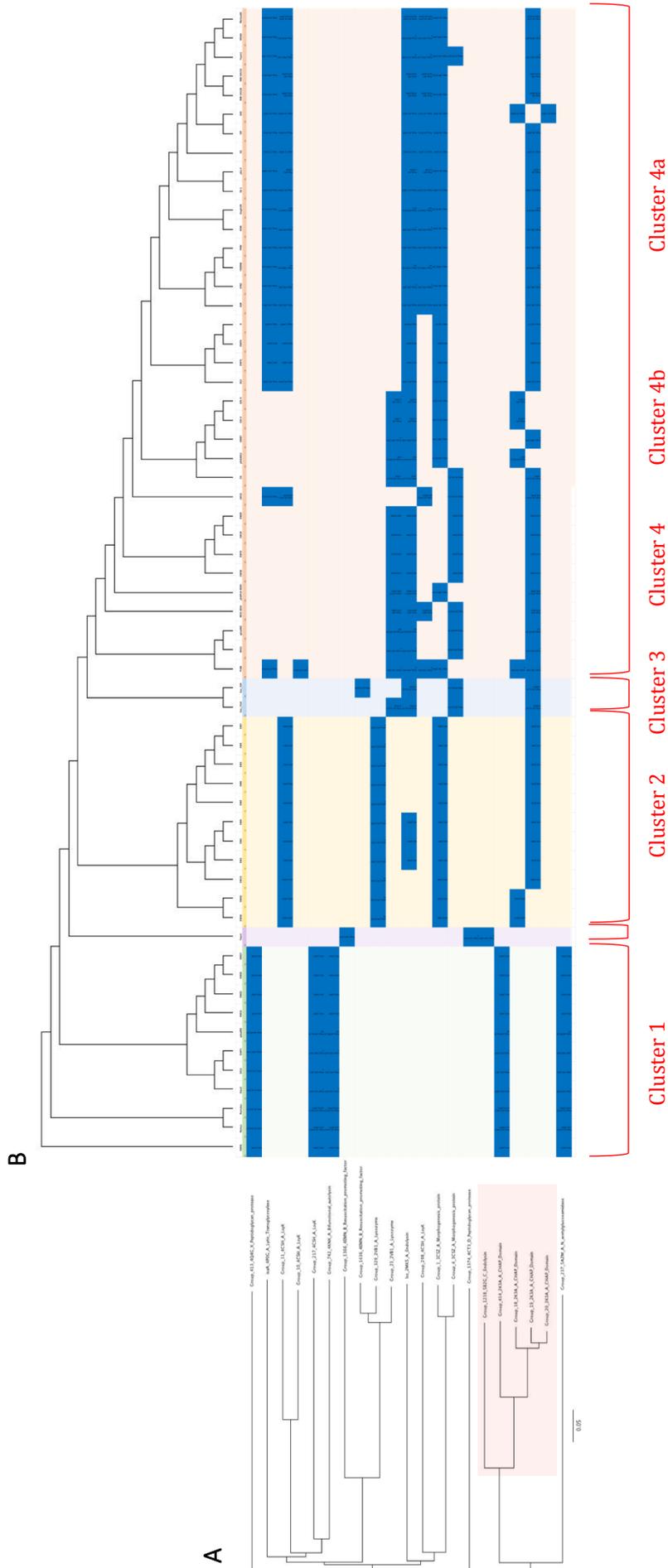
Figure 5.9: Gene presence/absence of predicted genes sharing structural homology to lytic enzymes among 60 *S. aureus* infecting bacteriophage in the subfamily *Twortvirinae*.

A) Dendrogram constructed using the single representative nucleotide sequence of each cluster (group) using Roary, each was translated in to its amino acid and aligned using MUSCLE with the

top hit for each lytic enzyme using HHpred alongside. B) Dendrogram of all 60 phage genomes built on Mash distances. Colours help distinguish unique phage clusters.

5.4 Discussion

We successfully sequenced the genomes of 22 study bacteriophage that displayed a broad host range against a large panel of clinical *Staphylococcus aureus* (including methicillin-resistant *S. aureus* – MRSA) isolates. The results of annotation and genomic comparison revealed that these phage are highly similar to members of the *Twortvirinae* subfamily. Due to the high degree of similarity of some of the study phage with other members of the cluster representing a distinct phage group, the unclassified study phage could therefore be included with that existing group. The genomes of 38 phage infecting *S. aureus* belonging to the *Twortvirinae* subfamily that were publicly available at the time were downloaded from the National Centre for Biotechnology Information (NCBI) and the European Nucleotide Archive (ENA).

Assembled genomes of the study phage were initially screened through BLASTn, all top hits aligned against *S. aureus* phage belonging to the *Twortvirinae* subfamily. We employed Mash to perform quantitative pairwise sequence comparisons between our study phage and all *Twortvirinae* infecting *S. aureus* that were publicly available using an all-versus-all approach, outlining several clusters and two singletons (Fig. 1). One of the major setbacks when working with large whole genome datasets is the significant computational power required to compute the rigorous nucleotide and amino acid sequence alignments using BLAST or Average Nucleotide Identity (ANI). Mash requires considerably less computational power enabling a much more rapid and efficient approach to estimate sequence similarity alignments based on kmers of assembled genome sequences by using the MiniHash technique for creating sketches to allow rapid distance estimates and resemblance between whole genomes without the need

for high memory requirements [243]. A collection of sequences are *hashed* in to MiniHash sketches and each sketch is compared between each other to estimate the fraction of shared k-mers (Jaccard index) and the rate of sequence mutation. By using the kmer-based approach it can be extremely useful when comparing highly diverse phage genomes. However, Mash calculates similarly without the requirement of alignment, therefore does not correlate with evolutionary divergence [334].

The output matrix was presented in a heatmap to demonstrate the pairwise comparison scores for each phage, all phage in this study displayed remarkably high homology between each other with the lowest being >68 %, a characteristic unique to *Spounavirinae* members [131]. Clusters were found to represent individual genera, with Cluster 1 made up of members belonging to *Silviavirus*, the single Twort virus and Cluster 4 consisting of *Kayvirus* being split into three smaller sub-clusters. Two smaller clusters and a singleton were also identified, all of which are still yet to be classified into a genus by the ICTV. Our study phage shared high similarity with previously sequenced phage, allocating them amongst the individual clusters. Additionally, cluster 2 consisting of only EW phage members sequenced in this study represents a distinct phage group, sharing only 91 % identity to the closest phage in Cluster 3, of which, members of this cluster are so far unclassified by the ICTV.

Mashtree provides a relatively simple but efficient approach to rapidly compare whole genome sequences in order to identify and study groups of similar phage that cluster together. Comparing predicted genes and the structural homology of the predicted protein those genes encode in combination with clustering is a relatively simple

approach as a means of identifying relationships between phage especially amongst phage within a specific genus. This approach has helped as to which family or genus the study phage may fall in to prior to morphological classification. The use of clustering has been used extensively when studying phage based on sequence identity and homology [338,351,352], however almost each one of these studies using a unique and different approach utilising an assortment of bioinformatics tools and programs. To appropriately classify the EW phage sequenced in this study, further genetic characterisation and morphological analysis will be required using ICTV approved methodology and programs, such as those that were utilised during the recent reorganisation of the current classification system for *Caudovirales* [131].

An initial pan-genome assessment comparing the predicted phage encoded proteins was performed in attempt to locate possible core genes conserved among all phage. However even when considering phage from the entire subfamily, no single gene was found to be shared across all phage genomes. As divergent as phages are, it would make sense that they should share genes that encode for key proteins essential to phage life. When examining the pan-genome, an immediately recognisable pattern are the blocks of genes arranged in rows shared across members unique to a cluster, closer examination of these clusters reveals that the members (of which were previously classified) also share the same genus. By comparing the genomes of phage within that individual cluster (genus), not only did they share a greater degree of similarity amongst each other, but we could identify core genes conserved among all its members. The identification of a core set of genes shared by a group of phage has been observed in similar studies, interestingly, they focused on just members of the same

187

genus, identifying a conserved core of essential genes including metabolic enzymes and structural components [352,353]. Phage lack the conserved, universal markers such as the 16S rRNA genes that form the basis of bacterial taxonomy, as there are no suitably conserved universal genes found in phage genomes [354]. With no single conserved protein marker shared across all phage, comparisons between phage diversity across genera and families is therefore very difficult. Consequently preventing a taxonomic system based on a single locus analogous to 16S rRNA approach for bacteria. Pairwise comparisons between sequences usually provide a robust method for species demarcation [355]. Based on genome similarity data obtained, several of the phage sequenced in this study appear to share indistinguishable levels of nucleotide identity with one another (100 %), suggesting that they are exactly the same phage. This could be down to various reasons, from the same phage being present among a number of tanks when collecting samples, to possible cross contamination during initial library preparation.

Surprisingly, the distribution of phage across genera is somewhat limited. Out of the 60 *Twortvirinae* genomes available, 35 (58 %) of those, belong to the *Kayvirus* genus and the remaining 27 (45 %) shared across the other genera. Interestingly, a distinctive characteristic of the *Kayvirus* is the relatively large accessory genome and the considerable gene diversity among those members, suggesting there are a number of hot spots for recombination. We note that the distribution of phage in clusters and the proportion of singletons (2 / 60 phage - 3.2 %), in addition to the two clusters composed of two individual members, is comparable with the proportion found in similar comparative phage genome studies, when investigating the diversity among

Enterobacteriaceae phage (18 / 338 – 5.3 %) [356], Mycobacteriophage (8 / 627 – 1.27 %) [341] and *Arthrobacter* phage (2 / 42 – 4.76 %) [357]. However, considerably lower compared to *Bacillus* phage (15 / 83 – 18 %) [358] and *Pseudomonas* phage (30 / 130 – 23.1 %) [342]. The gene coding potential for each of the phage genomes in this study averaged around 88 %, exhibiting a gene density of 1.47 genes per kilobase pair (kbp). These values are similar to that observed in other studies on *Pseudomonas aeruginosa* phage (93 % - 1.5 kbp) [339], *Bacillus* phage (88 % - 1.47 kbp) [358], *Arthrobacter* phage (1.52 kbp) [357], *Mycobacterio*phage (93 % - 1.42 kbp) [343] and a larger *Mycobacterio*phage study including members from the previous study (1.49 kbp) [341]. It must be noted that, these values were calculated by filtering out the *Twortvirinae* phage genome characteristics from tables and supplementary material provided.

The diversity of Twortvirinae was further demonstrated by the low level of genes shared among the clusters, as mentioned previously no single gene was present in all phage. The most commonly shared gene was found in 48 (80 %) phage out of the 60, belonging to clusters 2, 3 and 4 which shared a high degree of similarity among each other. Of the 1634 individual gene groups predicted by Roary, there were 583 (35.6 %) single gene groups found in one phage member, remarkably similar to those observed with mycobacteriophage (31 %) [341], and similarly with *Arthrobacter* phage (26 %) [357]. The proportions of cluster-associated core and accessory genes vary considerably depending on the clusters, demonstrating the level of diversity that is found between them. However, the encoded protein remains consistent across all clusters, as the majority of identified genes encode for essential proteins including major capsid, tail and baseplate proteins (structural), RNase, endonuclease and DNA

polymerase (metabolic), holin, endolysin and virion-associated peptidoglycan hydrolases (lytic enzymes).

With the rapidly increasing number of phage genomes being deposited in public databases, the possibility of using every single phage genome remains to be a problematic issue when attempting to perform comparative genomic studies. When this study first started in February 2018, only 38 genomes of phage infecting *S. aureus* were currently available at the time. However in the time it's taken to analyse this data, that number has almost doubled in size (http://millardlab.org/bioinformatics/ bacteriophage-genomes/). Furthermore, the number of phage infecting *S. aureus* currently available is rather difficult to identify. Not all publications on staphylococcal phage have reported their host species, and not every phage has a publication associated with it, therefore limiting the sample size of this study. Future studies will consider a more comprehensive dataset by including phage infecting an entire bacterial genus.

The varying quality and differences of annotated genomes across older and more recently deposited genomes, as well as multiple genome entries for a single phage genome can greatly affect gene production analysis [359]. To improve the original annotation for some of the genomes and ensuring uniformity, all 60 phage genomes used in this study were re-annotated with Prokka against a database of known viral proteins. Regardless of the database used, uncharacterised hypothetical proteins accounted for the majority of identified coding sequences, as is commonly the issue during annotation [344]. A considerable amount of information remains to be

elucidated from these phage, however, the need to enrich protein databases with identified genes remains a major challenge in attempt to successfully annotate phage genomes. The validity of this work significantly depends on the successful and correct identification of gene boundaries and coding sequences using updated protein databases, gene-predicting programs and manual annotation.

The presence of self-splicing elements such as group I introns have been previously described in tailed phage, especially within those infecting *S. aureus* [348]. Introns intervening gene sequences has become a challenging issue for annotation and gene prediction programs. For example, upon closer inspection of gene groups associated with a single phage, the presence of introns encoding for an endonuclease. These endonucleases are often found embedded within introns and appear to be a characteristic feature of *S. aureus* twortlikeviruses and found in a variety of genes such as, DNA polymerase, helicase and ribonucleotide reductase large-subunit [211,338,360]. Even with staphylococcal phage sharing the least similarity at the DNA level such as members of *Silviavirus* and *Kayvirus* genus, they still possess similar group I introns present within functional homologs [348].

As no single gene was found shared among all phage genomes, the predicted structure of genes were assessed and those that shared structural homology to previously identified lytic enzymes were screened. Most staphylococcal phage endolysins are multi-domain proteins harbouring a C-terminal SH3B-type cell wall binding domain (CBD), an amidase and a catalytic N-terminal endopeptidase domain (cysteine, histidine-dependant amidohydrolase/peptidase domain - CHAP) [361]. Phage

endolysins can be divided based on the target substrate in bacterial cell walls, their modular structure allows the possibility of different domains to be swapped producing different domain arrangements, thus enabling endolysins to have altered enzymatic specificity and efficacy [362]. We found that each of the phage clusters possessed a set of genes that were grouped in to a distinct branch, these genes all shared structural similarity to the same CHAP domain (2K3A_A), suggesting that the endolysins that each phage encoded most probably shared the same endopeptidase N-terminus. Notably, the 3-domain endolysin LysK, derived from the staphylococcal phage K that features a CHAP domain has been previously truncated to retain only the CHAP domain, has demonstrated a two-fold higher lytic activity against live *S. aureus* including MRSA [146].

Chapter 6

Discussion

6.1 Discussion

A key factor of *Staphylococcus aureus* virulence is its ability to form biofilms on a number of biotic and abiotic surfaces [363]. The production of biofilm is a common mechanism employed by bacteria as means of protection against the environment and components of the host immune system. Cells within biofilms can also be protected from the action of antibiotics and disinfectants leading to persistent, recalcitrant infections and delays in healing [364].

S. aureus (including methicillin-resistant *S. aureus* – MRSA) is a major cause of biofilm-associated infections frequently encountered within healthcare settings that can lead to disseminated, systemic infections that can result in patient death. Despite infection control measures that have reduced the numbers of MRSA outbreaks in UK hospitals [68], biofilm-associated *S. aureus* infections are still responsible for considerable rates of morbidity and mortality. Frequently multiply antibiotic resistant, these infections are usually treated with intravenous vancomycin to which intermediate susceptibility frequently emerges. Future strategies for the control of MRSA and other major antibiotic resistant pathogens necessitates the development of adjuncts and alternatives to antibiotics such as immunotherapy and bacteriophage therapy. These will hopefully offer future treatment options for infections that do not respond to antibiot*ic* chemotherapy.

The application of phage preparations in the treatment of bacterial infections has so far produced promising results. Lytic phage have been shown to be favourable candidates for therapeutic use and have already proven to be highly effective in the

control and treatment of *S. aureus* infections, including MRSA and their biofilms. The objective of this project was to assess the lytic potential of novel phage infecting *S. aureus* and determine their anti-biofilm properties against a number of major epidemic MRSA isolates belonging to two highly prevalent clonal lineages that are endemic in UK hospitals. To fulfil this aim, it was necessary to isolate and characterize a collection of phage from various environmental samples and laboratory stocks.

While phage are widespread within the environment, their presence correlates to the existence of their bacterial host and may only be present at low frequency. Difficulties in isolating lytic *S. aureus* bacteriophage from various sources including wastewater samples, faecal matter and environmental samples have been reported in numerous studies compared to enteric bacteria that are highly prevalent in sewage environments [267,365]. Lytic *S. aureus* phage are much less easy to isolate than those infecting other bacterial hosts such as the gram-negative species *Klebsiella pneumoniae, Pseudomonas aeruginosa, Escherichia coli* and *Salmonella* species [267,366,367].

Enrichment protocols can help create a bias towards the target phage with desirable characteristics for a specific host, where it can propagate and increase in number to detectable levels [367]. The enrichment method offers various opportunities to modify and tailor the protocol based on:- the type of environmental sample, the number and species of isolation hosts, detection of phage, formulation of growth media, the handling and processing of samples prior to infection and post infection [367]. The optimisation of a working enrichment protocol that would allow us to achieve the

greatest yields when attempting to isolate phage infecting *S. aureus* presented the first major hurdle in this study.

Initial sample enrichment attempts were performed using *S. aureus* D329 and H402 as hosts, resulting in low plaque yields. The high host specificity of phage greatly limits the likelihood of isolating possible phage present within the samples when using only two bacterial hosts and are otherwise missed. To counter this, one approach considered was to use more than a one host during sample enrichments by introducing a combination of several target hosts strains [290,368]. This bacterial cocktail approach has proven to be beneficial within previous studies isolating a number of phage effective against *S. aureus* [190,369]. However, increasing the diversity and numbers of bacterial hosts proved to be inefficient and biofilms produced by labour intensive, as replicates of each enrichment sample had to be plated out on to each of the hosts using the soft agar overlay during screening. The use of a modified *S. carnosus* TM300 strain which expresses N-acetyl-D-glucosamine-modified ribitol-phosphate (RboP) repeating units of the wall teichoic acid on to the surface of the cells, a known *S. aureus* phage receptor, was crucial to successfully isolating 46 novel obligately lytic phage from wastewater samples in this study. The addition of 32 phage from our laboratory stocks further expanded our collection to a total 78 phage as recorded in Chapter 3.

When considering phage as a potential therapeutic, obligately lytic phage are preferred due to their ability to rapidly kill the bacterial host, reducing the possibility of the host developing resistance. They should also lack genes responsible for genome incorporation that may mediate horizontal gene transfer of virulence or antibiotic

resistance genes. Determination of the host range (lytic spectrum) for each newly isolated phage should be assessed against a host panel in order to determine the level of polyvalence. We tested each phage against a genetically diverse representative group of 185 *S. aureus* isolates including a number of major MRSA clones. Host range analysis demonstrated that 40 of the 78 phage were able to infect over 90 % of isolates tested. The broad host range exhibited among the large majority of our phage collection shared similar broad host range properties to notable polyvalent phage that belong to two of the common and widely used (in phage therapy) staphylococcal phage groups, the Twort-Like (TL) myoviruses and lytic podoviruses (LP) [202].

One issue to consider in selecting phage isolates for therapeutic development is the minimum number of susceptible bacterial hosts strains required in testing to adequately determine whether a phage exhibits a broad or narrow host range. This was recently addressed by a number of phage scientists [367], feedback from this showed very little agreement on the size of the panel used which ranged from single figures up to 800 in some cases. Several respondents indicated that there were multiple variables to consider depending on the bacterial host, such as the diversity among test isolates and the proportion of closely- or distantly- related host of that selected species. Additionally, the planned use of the phage will also be important to consider, for example whether it will used to specifically target strains isolated from clinical, environmental or animal sources. Studies rarely characterise phage against the same bacterial isolates and panel size as others making it difficult to assess phage host range from study to study. One consideration could be the introduction of an international standard panel unique to the target host species or even genus to be

included for every test, consisting of well-studied 'type strains' host or members from highly successful and prevalent lineages associated with epidemics.

Killing assays were employed to further assess the efficacy of each phage against planktonic cultures being demonstrated *in vitro*. Ten phage from our collection that previously exhibited a relatively broad but varied host range with coverage of almost all bacterial isolates within the panel tested against were selected for this experiment. The treatment of three *S. aureus* isolates and a *S. carnosus* TM300 host used for phage isolation provided a detailed representation of the unique growth kinetics of each phage whilst highlighting the variation in host specificity in liquid culture. This assay monitored the bactericidal ability of each phage in real-time by measuring changes in absorbance, whilst being able to observe the growth of phage resistant mutants.

Differences in virulence factors among *S. aureus* isolates has been linked to specific lineages based on sequence type (ST) [72,370] or *spa* protein [371], with some clones demonstrating a higher propensity to form stronger biofilms or greater adherence to surfaces than others [224,372]. The association between biofilm production capacity and clonal lineages[224,372], specifically by ST22 and ST36 clones was assessed in Chapter 4. Due to the relatively low biofilm levels produced by our collection of ST22 isolates, a cut-off value of 0.4 (OD$_{590}$) was used to distinguish between strong and weak biofilm formers, resulting in only seven isolates being considered as strong biofilm formers. Biofilm biomass was significantly greater among ST36 isolates with two isolates in particular capable of producing three times more biofilm mass than the highest producing ST22 isolates. To allow a more accurate comparison, two of the

weaker biofilm producers of ST36 that achieved densities similar to those obtained by the most successful biofilm producers of ST22 were selected. The selected isolates used in this study demonstrated an association between biofilm forming capacity and clonal lineage, in agreement with those observed in previous *S. aureus* biofilm studies [227,363].

However, when considering the most successful biofilm formers, our results were not consistent with those revealed in a previous study, out of the 13 clonal complexes studied the EMRSA-15 clone (ST22) appeared to be the more successful at forming stronger biofilms when compared to EMRSA-16 isolates (ST36) [363]. It must be noted that in this study the authors used lower starting inoculum diluted to an optical density (600 nm) of 0.04 to grow their biofilms, whilst also using the Calgary biofilm device (polystyrene pegs on the coverlid) approach. This allows biofilms to form on the pegs of the lid as a resultant of sessile development rather than cell sediment. Therefore, it is important to consider that differences in biofilm strength and formation between studies can be largely associated with the growth conditions and protocol used, that may include the various type of plates, different media strengths, supplemented with glucose used for the assay [373].

In a clinical scenario, treatment of severe *S. aureus* infections is usually performed using a combination of antibiotics to further enhance their activity in attempt to reduce and prevent and reduce the emergence of antibiotic resistant mutants [374]. However, bacteria embedded within biofilms are inherently difficult to treat with antimicrobials and bactericidal agents, demonstrating a reduced susceptibility to antibiotics. The anti-

biofilm properties of phage and their enzymes make them an attractive alternative to antibiotics and an effective antimicrobial towards controlling bacterial population dynamics within *S. aureus* biofilms. The glycocalyx and EPS may hinder the diffusion of phage particles, however the concentration of lytic enzymes and peptidoglycan hydrolase activity may play a role in the degradation and disruption of the biofilm matrix, thus facilitating phage diffusion enabling phage to reach cells that were once protected. To test this statement, the four most effective phage based on their success in Chapter 3, were selected to further assess the efficacy and anti-biofilm capacity against four ST22 and four ST36 clinical isolates using the microtitre plate biofilm assay. In this study, each of the four phage demonstrated high efficacy against biofilms and were successful in greatly reducing cell viability and biofilm biomass following single application for 6 and 24 h at an MOI 1 and 0.1.

Overall, reductions in biofilm biomass were generally greater in biofilms when treated with an MOI 0.1 for each of the phage. Interestingly, we found that biofilm densities did not correlate with viable cells numbers, biofilm cell regrowth was observed when exposed to phage for 24 h than 6, however resistance may have come at the expense of biofilm production as no further increase in biofilm densities was observed during both sampling timepoints. The inability to completely eradicate infection-causing pathogens within biofilms and development of phage resistance, presents a major limitation when considering the therapeutic potential of phage. Previous studies have demonstrated the bacteria and phage undergo antagonistic coevolution *in vitro.* Suggesting that over time, bacteria develop and become more resistant to infective phage when exposed for long periods of time. However as time progresses, so too does

the phage, subsequently evolving to become more infectious to the resistant bacteria [375].

However, the ability of phage to successfully disperse biofilms but not completely kill all bacteria can still work as an advantage, the prospect that phage could be used in conjunction with antibiotics and other antimicrobials agents in the treatment of biofilms has been previously considered and extensively reviewed [376,377]. By utilising phage and their enzymes to disrupt the key components of the biofilm EPS, it exposes the once protected cells that have been found to be susceptible to sub-inhibitory concentrations of antibiotics in attempt to prolong their utility. Studies have identified potential synergistic relationships between phage and antibiotics, the combined or sequential administration of both of these has been shown to improve their efficacy, resulting in enhanced biofilm disruption and reduction to viable cells *in vitro* [217,219].

Historically, morphology-based classification was the main approach used as a means of characterising and gaining a better understanding of phage taxonomy. One of the mandatory taxonomical classification criteria for categorising phage is the examination of its morphology characteristics through electron microscopy, thus allowing to determine the family based on tail morphology and its nucleic acid. However, the significant increase in numbers of novel phage genomes being sequenced and submitted to public databases every year has revealed that tailed-phage are much more diverse than originally thought [131]. Alternative and more modern approaches for genomically classifying phage have been considered as many biologists feel the

current taxonomy is outdated [109]. Chapter 5 describes the whole-genome sequencing of 22 novel lytic phage genomes. These exhibit high sequence homology to previously characterised phage infecting *S. aureus* belonging to the newly established *Twortvirinae* subfamily. The complete genomes of all 38 phage infecting *S. aureus* were download from public databases and comparative genomic analyses was subsequently carried out using Mash distances. This resolved several clusters with phage from this study grouping with clusters containing previously described phage genomes. However a small number of phage isolated in this study formed a distinct cluster to that of other phage genomes. Further assessment revealed each unique cluster formed using MASH distances represents individual genera within the *Twortvirinae* subfamily with members within each cluster sharing a high level of homology between each other. The assignment of phage genomes in to clusters closely reflects the relationship among phage groups based on sequence similarity or gene content and has been employed extensively when performing comparative analyses with large genome datasets [341,351,358]. This type of clustering based on pairwise distances does not substitute for a reticulate taxonomy, however it may be employed to extend the classification of staphylococcal phage, providing insight in to the assignment of phage in to families or possible genera prior to morphological classification.

Further analysis of selected phage genomes in Chapter 5 identified a number of phage from this study that demonstrated indistinguishable levels of identity between one another and were found to be identical. The cause for this could be down to a number of possible factors, suggesting that these phage were most probably present in a

number of ASP sewage tanks when samples were collected, possible contamination during the processing of wastewater samples during screening, or operator error during phage when isolating and purifying phage plaques. Pan-genome analysis allowed for gene presence/absence comparisons amongst the study group, revealing no single gene was shared between all phage, with the most commonly shared gene present among 80 % (48/60) of the phage collection. However, lack of gene content present between phage made comparative genomics of the phage collection challenging. Furthermore a significant proportion of the gene groups identified (35.6 %) were only present in a single phage, yet further interrogation of these gene sequences revealed the presence of Group I introns that often encode for homing endonucleases inserted within key metabolic and structural genes that were classed as individual gene groups but were otherwise identical [348,378]. By utilising HHpred, the majority of genes that were shared among phage shared structural protein homology with lytic enzymes and virion structural proteins. Genes encoding for lytic enzymes were further assessed, revealing all phage encoded for a protein that shared high structural similarity to the same CHAP domain protein, suggesting that each phage encodes an endolysin that most probably shared the same endopeptidase N-terminus.

The number of phage genome sequences being deposited and made available on public databases continues to rapidly increase, with staphylococcal phage genomes almost doubling in size since this study started. One issue was the number of sequenced genomes of lytic phage infecting *S. aureus* that were currently available on public databases was difficult to determine. Of course, not every phage has a publication associated with it that will aid in determining the bacterial host of the

phage, however the lack of supporting detail and information provided for each phage sequence when initially submitted greatly limited the sample size we used. An informal nomenclatural guide on how to name and classify phage was recently introduced in 2017, proposing each novel phage to be uniformly named according to unique identifiers such as bacterial host genus and phage morphology to be provided, whilst helping to eliminate the existing issue of having multiple diversely-related phage possessing the same name [275].

6.2 Suggestions for future work

Further characterisation of the wide-host range phage from this study should be investigated to develop a greater understanding of their biology. This would include experiments to determine the parameters of growth kinetics such as adsorption rate, burst size, eclipse period and lysis time. Methods could also be used to extend the host range of those exhibiting a narrow spectrum perhaps using a directed evolution approach [379]. Additionally, morphological classification using transmission electron microscopy (TEM) continues to be one of the major criterion for phage classification and considered to confirm the phage morphology derived by genome similarities.

Four phage, isolated in this study, were highly efficacious in reducing biofilm mass and number of viable cells of *S. aureus* ST22 and ST36 isolates in static microtitre plate assays. However this static biofilm system is a relatively crude model - further assessment of these phage should be conducted using a dynamic biofilm model with continuous flow of nutrients which generates a more realistic (and clinically relevant) environment for biofilms to develop in. Such a system using, for example, a modified

Robbins device or a Drip Flow Biofilm Reactor could be used to generate robust datasets to describe the *in vitro* characteristics of our phage against *S. aureus* biofilms. It would also be beneficial to assess the action of phage to manipulate biofilms grown on to range of medical surfaces such as those used for catheters and implants should also be assessed due to the significant rate of implant associated infections associated with *S. aureus*.

Nevertheless, phage virulence *in vitro* does not always correlate with virulence and efficacy *in vivo* so evaluation of phage efficacy *in vivo* using using animal models such as the mouse or rat would be required as part of a clinical development pathway. As *in vivo* environments are highly complex and dynamic, bacterial behaviour can significantly differ from that of the *in vitro* state which will impact on effect on phage kinetics. Overall gene expression can also change, leading to loss of the surface receptors phage require to bind to, which may also lead to reduced sensitivity to phage [380–382]. These interactions further increase the complexity of model, greatly restricting the ability to screen multiple hosts and phage, therefore the choice of animal model should be carefully considered. Furthermore, it is crucial to ensure the purity and safety of phage lysates and preparations prior to administration of phage in animal systems.

The sequencing of the remaining 56 phage within our phage collection will help to further enrich the current databases with more staphylococcal phage genomes, this is important as genome sequence databases are accessed at a much higher rate than a traditional publication. This will allow other research scientists to access a large

collection of phage genomes to help in efforts to explore phage genomic diversity and evolution and understand the complexity of phage genomes and their relation to that of their host. Undoubtedly more *Twortvirinae* will be discovered in the future and their sequenced genomes will be uploaded and made available in the public domain. More work should be invested in to studying the phages that are currently unclassified, although undersampled, they may represent a new genus or possibly subfamily, and it is by enriching these databases that will help to assign these phage in to their respective genus or subgroups. To consider genome wide association studies (GWAS) and pan-GWAS to further investigate the biofilm production between ST22 and ST36 isolates. Using pan-genome pipelines such as Roary it is possible to identify candidate genes that are present or absent within good and weak biofilm producers, and assess their contribution to biofilm formation [383].

The high mutability of *S. aureus* enables the pathogen to rapidly adapt when exposed to external stressors. It would also be interesting to study the phage-resistant mutants producing various morphotypes that were recovered from phage treated biofilms. Whole genome sequencing of the various mutants and untreated isolates would enable genomic comparisons to look for genomic changes such as mutations, deletions and insertions within the genomes. This would provide a greater understanding of the adaptive strategy employed by biofilm-associated *S. aureus*.

6.3 Concluding remarks

In summary, this study attempted to assess the lytic activity of phage infecting *S. aureus*. By utilising a modified *S. carnosus* TM300 with an increase susceptibility to

phage adsorption, we successfully isolated 46 novel obligately lytic phage from wastewater samples that exhibited a broad host range determined by testing against an international panel of genetically-characterised MRSA isolates. With emphasis on the phage-biofilm interaction of candidate phage against mature biofilms produced by two highly successful and dominant MRSA clones within the UK, ST22 and ST36. The top four candidates demonstrated promising anti-biofilm activity against strong biofilm producers, however reductions in biofilm biomass did not correlate with reductions in viable cells, as phage failed to completely eradicate all adhered cells. Genomic comparisons of 22 study phage identified them to share significant nucleotide similarity to *Twortvirinae* phage previously classified, no core genome was identified, they did all carry a gene that shared significant structural homology to the same CHAP domain. Overall, the phage isolated in this study may be useful therapeutic candidates when applied in combination with other phage or antibiotics in attempt to treat MRSA infections in the future.

References

1. Ventola CL. The Antibiotic Resistance: part 1: causes and threats. P&T. 2015;40:277–83.

2. Cassini A, Högberg LD, Plachouras D, Quattrocchi A, Hoxha A, Simonsen GS, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis. 2019;19:56–66.

3. O 'neill J. Tackling Drug-Resistant Infections Globally: Final Report and Recommendations the Review on Antimicrobial Resistance. 2016;1–84. Available from: https://amr-review.org/sites/default/files/160525_Final paper_with cover.pdf

4. Fair RJ, Tor Y. Antibiotics and bacterial resistance in the 21st century. Perspect Medicin Chem. 2014;6:25–64.

5. Campbell A. The future of bacteriophage biology. Nat Rev Genet. 2003;4:471–7.

6. Lowy FD. *Staphylococcus aureus* infections. N Engl J Med. 1998;339:520–32.

7. Dinges MM, Orwin PM, Schlievert PM. Exotoxins of *Staphylococcus aureus*. Clin Microbiol Rev. 2000;13:16–34.

8. Kluytmans J, Van Belkum A, Verbrugh H, Kluytmans J, van Belkum A VH. Nasal carriage of *Staphylococcus aureus*: Epidemiology, underlying mechanisms, and associated risks. Clin Microbiol Rev. 1997;10:505–20.

9. Foster TJ. The *Staphylococcus aureus* "superbug." J Clin Invest. 2004;114:1693–6.

10. van Belkum A, Verkaik NJ, de Vogel CP, Boelens HA, Verveer J, Nouwen JL, et al. Reclassification of *Staphylococcus aureus* nasal carriage types. J Infect Dis. 2009;199:1820–6.

11. Gordon RJ, Lowy FD. Pathogenesis of Methicillin-Resistant *Staphylococcus aureus* Infection. Clin Infect Dis. 2008;46.

12. Brandt CM, Duffy MCT, Berbari EF, Hanssen AD, Steckelberg JM, Osmon DR.

*Staphylococcus aureus*s prosthetic joint infection treated with prosthesis removal and delayed reimplantation arthroplasty. Mayo Clin Proc. 1999;74:553–8.

13. de Lissovoy G, Fraeman K, Hutchins V, Murphy D, Song D, Vaughn BB. Surgical site infection: Incidence and impact on hospital utilization and treatment costs. Am J Infect Control. 2009;37:387–97.

14. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005;3:722–32.

15. Malachowa N, Deleo FR. Mobile genetic elements of *Staphylococcus aureus*. Cell Mol Life Sci. 2010;67:3057–71.

16. Herron-Olson L, Fitzgerald JR, Musser JM, Kapur V. Molecular correlates of host specialization in *Staphylococcus aureus*. PLoS One. 2007;2:e1120.

17. Arvidson S, Tegmark K. Regulation of virulence determinants in *Staphylococcus aureus*. Int J Med Microbiol. 2001;291:159–70.

18. Patti JM, Allen BL, McGavin MJ, Hook M, Höök M. MSCRAMM-mediated adherence of microorganisms to host tissues. Annu Rev Microbiol. 1994;48:585–617.

19. Edwards AM, Potts JR, Josefsson E, Massey RC. *Staphylococcus aureus* host cell invasion and virulence in sepsis is facilitated by the multiple repeats within FnBPA. PLoS Pathog. 2010;6:e1000964.

20. Kobayashi SD, DeLeo FR. *Staphylococcus aureus* protein A promotes immune suppression. MBio. 2013;4:e00764-13.

21. Athanasopoulos AN, Economopoulou M, Orlova V V., Sobke A, Schneider D, Weber H, et al. The extracellular adherence protein (Eap) of *Staphylococcus aureus* inhibits wound healing by interfering with host defense and repair mechanisms. Blood. 2006;107:2720–7.

22. Foster TJ. The remarkably multifunctional fibronectin binding proteins of *Staphylococcus aureus*. Eur J Clin Microbiol Infect Dis. 2016;35:1923–31.

23. Kong C, Neoh HM, Nathan S. Targeting *Staphylococcus aureus* toxins: A potential form of anti-virulence therapy. Toxins (Basel). 2016;8:72.

24. Foster TJ. Immune evasion by staphylococci. Nat Rev Microbiol. 2005;3:948–58.

25. Shallcross LJ, Fragaszy E, Johnson AM, Hayward AC. The role of the Panton-Valentine leucocidin toxin in staphylococcal disease: A systematic review and meta-analysis. Lancet Infect Dis. 2013;13:43–54.

26. Decker CF. Pathogenesis of MRSA Infections. Disease-a-Month. 2008;12:774–9.

27. Brussow H, Canchaya C, Hardt W-D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiol Mol Biol Rev. 2004;68:560–602.

28. Gillet Y, Issartel B, Vanhems P, Fournet JC, Lina G, Bes M, et al. Association between Staphylococcus aureus strains carrying gene for Panton-Valentine leukocidin and highly lethal necrotising pneumonia in young immunocompetent patients. Lancet. 2002;359:753–9.

29. Löffler B, Hussain M, Grundmeier M, Brück M, Holzinger D, Varga G, et al. *Staphylococcus aureus* panton-valentine leukocidin is a very potent cytotoxic factor for human neutrophils. PLoS Pathog. 2010;6.

30. Wang JT, Wang JL, Fang CT, Chie WC, Lai MS, Lauderdale TL, et al. Risk factors for mortality of nosocomial methicillin-resistant *Staphylococcus aureus* (MRSA) bloodstream infection: With investigation of the potential role of community-associated MRSA strains. J Infect. Elsevier Ltd; 2010;61:449–57.

31. Mishra AK, Yadav P, Mishra A. A systemic review on Staphylococcal scalded skin syndrome (SSSS): A rare and critical disease of neonates. Open Microbiol J. 2016;10:150.

32. McCormick JK, Yarwood JM, Schlievert PM. Toxic shock syndrome and bacterial superantigens: an update. Annu Rev Microbiol. 2002;

33. Kulhankova K, Kinney KJ, Stach JM, Gourronc FA, Grumbach IM, Klingelhutz AJ, et al. The superantigen toxic shock syndrome toxin 1 alters human aortic endothelial cell function. Infect Immun. 2018;55:77–104.

34. Willems HM, Xu Z, Peters BM. Polymicrobial biofilm studies: from basic science to biofilm Ccontrol. Curr Oral Heal Reports. 2016;3:36–44.

35. Gabrilska RA, Rumbaugh KP. Biofilm models of polymicrobial infection. Future Microbiol. 2015;10:1997–2015.

36. Flemming HC, Wingender J, Allison DG. The biofilm matrix. Nat Rev Microbiol. Nature Publishing Group; 2010;8:623–33.

37. Costerton JW. Bacterial biofilms: a common cause of persistent infections. Science (80-). 1999;284:1318–22.

38. Cerca N, Oliveira R, Azeredo J. Susceptibility of Staphylococcus epidermis planktonic cells and biofilms to the lytic action of *Staphylococcus* bacteriophage K. Lett Appl Microbiol. 2007;45:313–7.

39. Wood TK. Combatting bacterial persister cells. Biotechnol Bioeng. 2016;113:476–83.

40. Tuomanen E, Durack DT, Tomasz A. Antibiotic tolerance among clinical isolates of bacteria. Antimicrob Agents Chemother. 1986;30:521.

41. Høiby N, Bjarnsholt T, Givskov M, Molin S, Ciofu O. Antibiotic resistance of bacterial biofilms. Int J Antimicrob Agents. 2010;35:322–32.

42. Arciola CR, Campoccia D, Speziale P, Montanaro L, Costerton JW. Biofilm formation in Staphylococcus implant infections. A review of molecular mechanisms and implications for biofilm-resistant materials. Biomaterials. 2012;33:5967–82.

43. Paharik AE, Horswill AR. The staphylococcal biofilm: Adhesins, regulation, and host response. Virulence Mech Bact Pathog Fifth Ed. 2016. p. 529–66.

44. Francois P, Schrenzel J, Stoerman-Chopard C, Favre H, Herrmann M, Foster TJ, et al. Identification of plasma proteins adsorbed on hemodialysis tubing that promote *Staphylococcus aureus* adhesion. J Lab Clin Med. 2000;135:32–42.

45. Monds RD, O'Toole GA. The developmental model of microbial biofilms: ten years of a paradigm up for review. Trends Microbiol. 2009;17:73–87.

46. O'Gara JP. ica and beyond: Biofilm mechanisms and regulation in *Staphylococcus epidermis* and *Staphylococcus aureus*. FEMS Microbiol Lett. 2007;270:179–88.

47. Moormeier DE, Bayles KW. *Staphylococcus aureus* biofilm: a complex

developmental organism. Mol Microbiol. 2017;104:365–76.

48. Otto M. Staphylococcal Infections: Mechanisms of biofilm maturation and detachment as critical determinants of pathogenicity. Annu Rev Med. 2012;64:175–88.

49. Dunne WM. Bacterial adhesion: seen any good biofilms lately? Clin Microbiol Rev. 2002;270:179–88.

50. Stoodley P, Sauer K, Davies DG, Costerton JW. Biofilms as complex differentiated communities. Annu Rev Microbiol. 2002;56:187–209.

51. Fleming D, Rumbaugh K. The consequences of biofilm dispersal on the host. Sci Rep. 2018;8:1–7.

52. Yarwood JM, Bartels DJ, Volper EM, Greenberg EP. Quorum sensing in *Staphylococcus aureus* biofilms. J Bacteriol. 2004;186:1838–50.

53. Donlan RM. Biofilms and device-associated infections. Emerg Infect Dis. 2001;7:277–81.

54. James GA, Swogger E, Wolcott R, Pulcini ED, Secor P, Sestrich J, et al. Biofilms in chronic wounds. Wound Repair Regen. 2008;16:37–44.

55. Baldoni D, Haschke M, Rajacic Z, Zimmerli W, Trampuz A. Linezolid alone or combined with rifampin against methicillin-resistant *Staphylococcus aureus* in experimental foreign-body infection. Antimicrob Agents Chemother. 2009;53:1142–8.

56. Lowy FD. *Staphylococcus aureus* Infections. N Engl J Med. 1998;

57. Baldan R, Testa F, Lorè NI, Bragonzi A, Cichero P, Ossi C, et al. Factors contributing to epidemic MRSA clones replacement in a hospital setting. PLoS One. 2012;7:e43153.

58. Lindsay JA, Moore CE, Day NP, Peacock SJ, Witney AA, Stabler RA, et al. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J Bacteriol. 2006;188:669–76.

59. Peacock SJ, Moore CE, Justice A, Kantzanou M, Story L, Mackie K, et al. Virulent combinations of adhesin and toxin genes in natural populations of *Staphylococcus*

aureus. Infect Immun. 2002;70:4987–98.

60. Van Belkum A, Melles DC, Snijders S V., Van Leeuwen WB, Wertheim HFL, Nouwen JL, et al. Clonal distribution and differential occurrence of the enterotoxin gene cluster, egc, in carriage- versus bacteremia-associated isolates of *Staphylococcus aureus*. J Clin Microbiol. 2006;44:1555–7.

61. Ben Zakour NL, Sturdevant DE, Even S, Guinane CM, Barbey C, Alves PD, et al. Genome-wide analysis of ruminant *Staphylococcus aureus* reveals diversification of the core genome. J Bacteriol. 2008;190:6302–17.

62. Lindsay JA, Holden MTG. Understanding the rise of the superbug: Investigation of the evolution and genomic variation of *Staphylococcus aureus*s. Funct Integr Genomics. 2006;6:186–201.

63. Novick RP, Subedi A. The SaPIs: Mobile pathogenicity islands of *Staphylococcus*. Chem Immunol Allergy. 2007;93:42–57.

64. Green BN, Johnson CD, Egan JT, Rosenthal M, Griffith EA, Evans MW. Methicillin-resistant*Staphylococcus aureus*: An overview for manual therapists. J Chiropr Med. 2012;11:64–76.

65. Goerke C, Pantucek R, Holtfreter S, Schulte B, Zink M, Grumann D, et al. Diversity of prophages in dominant *Staphylococcus aureus*clonal lineages. J Bacteriol. 2009;191:3462–8.

66. Lindsay JA. Genomic variation and evolution of *Staphylococcus aureus*. Int J Med Microbiol. Elsevier; 2010;300:98–103.

67. Jevons MP, Rolinson GN, Knox R. Celbenin-resistant staphylocci. Br Med J. 1961;1:124.

68. Andersen BM. MRSA Prevention. Prev. Control Infect. Hosp. Pract. Theory. 2019.

69. Hiramatsu K, Ito T, Tsubakishita S, Sasaki T, Takeuchi F, Morimoto Y, et al. Genomic basis for methicillin resistance in *Staphylococcus aureus*. Infect Chemother. 2013;45:117–36.

70. Francis JS, Doherty MC, Lopatin U, Johnston CP, Sinha G, Ross T, et al. Severe community-onset pneumonia in healthy adults caused by methicillin-resistant

*Staphylococcus aureus*carrying the panton-valentine leukocidin genes. Clin Infect Dis. 2004;40:100–7.

71. Gonzalez BE, Hulten KG, Dishop MK, Lamberth LB, Hammerman WA, Mason EO, et al. Pulmonary manifestations in children with invasive community-acquired *Staphylococcus aureus* infection. Clin Infect Dis. 2005;41:583–90.

72. Deurenberg RH, Stobberingh EE. The evolution of *Staphylococcus aureus*. Infect Genet Evol. 2008;8:747–63.

73. Hidron AI, Low CE, Honig EG, Blumberg HM. Emergence of community-acquired meticillin-resistant *Staphylococcus aureus* strain USA300 as a cause of necrotising community-onset pneumonia. Lancet Infect Dis. 2009;9:384–92.

74. David MZ, Cadilla A, Boyle-Vavra S, Daum RS. Replacement of HA-MRSA by CA-MRSA infections at an academic medical center in the midwestern United States, 2004-5 to 2008. PLoS One. 2014;9:e92760.

75. Song JH, Hsueh PR, Chung DR, Ko KS, Kang CI, Peck KR, et al. Spread of methicillin-resistant *Staphylococcus aureus* between the community and the hospitals in Asian countries: An ANSORP study. J Antimicrob Chemother. 2011;66:1061–9.

76. Enright MC, Day NPJ, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. J Clin Microbiol. 2000;38:1008–15.

77. Feil EJ, Cooper JE, Grundmann H, Robinson DA, Enright MC, Berendt T, et al. How clonal is *Staphylococcus aureus*? J Bacteriol. 2003;185:3307–16.

78. Abdelbary MMH, Basset P, Blanc DS, Feil EJ. The evolution and dynamics of methicillin-resistant *Staphylococcus aureus*. Genet Evol Infect Dis Second Ed [Internet]. Elsevier Inc.; 2017. p. 553–72. Available from: http://dx.doi.org/10.1016/B978-0-12-799942-5.00024-X

79. David MZ, Daum RS. Community-associated methicillin-resistant *Staphylococcus aureus*: Epidemiology and clinical consequences of an emerging epidemic. Clin Microbiol Rev. 2010;23:616–87.

80. Enright MC, Robinson DA, Randle G, Feil EJ, Grundmann H, Spratt BG. The

evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA). Proc Natl Acad Sci [Internet]. 2002;99:7687–92. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.122108599

81. DeLeo FR, Otto M, Kreiswirth BNB, Chambers HFH. Community-associated meticillin-resistant *Staphylococcus aureus*. Lancet [Internet]. 2010;375:1557–68. Available from: http://www.sciencedirect.com/science/article/pii/S0140673609619991%5Cnhttp://dx.doi.org/10.1016/S0140-6736(09)61999-1

82. Kim J. Understanding the evolution of methicillin-resistant *Staphylococcus aureus*. Clin Microbiol Newsl [Internet]. Elsevier Inc.; 2009;31:17–23. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0196439909000038

83. Chambers HF, DeLeo FR. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. Nat Rev Microbiol [Internet]. Nature Publishing Group; 2009;7:629–41. Available from: http://dx.doi.org/10.1038/nrmicro2200

84. Stenhem M, Örtqvist Å, Ringberg H, Larsson L, Olsson-Liljequist B, Hæggman S, et al. Imported methicillin-resistant Staphylococcus aureus, Sweden. Emerg Infect Dis. 2010;16:189.

85. Hassoun A, Linden PK, Friedman B. Incidence, prevalence, and management of MRSA bacteremia across patient populations-a review of recent developments in MRSA management and treatment. Crit Care. Critical Care; 2017;21:211.

86. Robinson DA, Enright MC. Evolutionary models of the emergence of methicillin-resistant *Staphylococcus aureus*. AntimicrobAgents Chemother [Internet]. 2003;47:3926–34. Available from: c:%5CKarsten%5CPDFs%5CStaphylokokken-PDFs%5CStaph-2003%5CRobinson - Enright-Evolutionary models of the emergence of methicillin-resistant S.aureus.pdf

87. Moore PCL, Lindsay JA. Molecular characterisation of the dominant UK methicillin-resistant *Staphylococcus aureus* strains, EMRSA-15 and EMRSA-16. J Med Microbiol. 2002;51:516–21.

88. Cox RA, Conquest C, Mallaghan C, Marples RR. A major outbreak of methicillin-

resistant *Staphylococcus aureus* caused by a new phage-type (EMRSA-16). J Hosp Infect. 1995;29:87–106.

89. Murchan S, Aucken HM, O'Neill GL, Ganner M, Cookson BD. Emergence, spread, and characterization of phage variants of epidemic methicillin-resistant *Staphylococcus aureus* 16 in England and Wales. J Clin Microbiol. 2004;42:5154–60.

90. McAdam PR, Templeton KE, Edwards GF, Holden MTG, Feil EJ, Aanensen DM, et al. Molecular tracing of the emergence, adaptation, and transmission of hospital-associated methicillin-resistant *Staphylococcus aureus*. Proc Natl Acad Sci [Internet]. 2012;109:9107–12. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1202869109

91. Ellington MJ, Hope R, Livermore DM, Kearns AM, Henderson K, Cookson BD, et al. Decline of EMRSA-16 amongst methicillin-resistant *Staphylococcus aureus* causing bacteraemias in the UK between 2001 and 2007. J Antimicrob Chemother. 2009;65:446–8.

92. Knight GM, Budd EL, Whitney L, Thornley A, Al-Ghusein H, Planche T, et al. Shift in dominant hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) clones over time. J Antimicrob Chemother. 2012;67:2514–22.

93. Richardson JF, Reith S. Characterization of a strain of methicillin-resistant *Staphylococcus aureus* (EMRSA-15) by conventional and molecular methods. J Hosp Infect. 1993;25:45–52.

94. Melter O, Urbásková P, Jakubů V, Macková B, Zemlicková H. Emergence of EMRSA-15 clone in hospitals throughout the Czech Republic. Euro Surveill. 2006;11.

95. Grundmann H, Aanensen DM, Van Den Wijngaard CC, Spratt BG, Harmsen D, Friedrich AW, et al. Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: A molecular-epidemiological analysis. PLoS Med. 2010;7.

96. Witte W, Enright M, Schmitz FJ, Cuny C, Braulke C, Heuck D. Characteristics of a new epidemic MRSA in Germany ancestral to United Kingdom EMRSA 15. Int J Med Microbiol. 2001;290:677–82.

97. Amorim ML, Faria NA, Oliveira DC, Vasconcelos C, Cabeda JC, Mendes AC, et al.

Changes in the clonal nature and antibiotic resistance profiles of methicillin-resistant *Staphylococcus aureus* isolates associated with spread of the EMRSA-15 clone in a tertiary care Portuguese hospital. J Clin Microbiol. 2007;45:2881–8.

98. Nadig S, Raju SR, Arakere G. Epidemic meticillin-resistant *Staphylococcus aureus* (EMRSA-15) variants detected in healthy and diseased individuals in India. J Med Microbiol. 2010;59:815–21.

99. Hsu LY, Loomba-Chlebicka N, Koh YL, Tan TY, Krishnan P, Lin RTP, et al. Evolving EMRSA-15 epidemic in Singapore hospitals. J Med Microbiol. 2007;56:376–9.

100. Vindel A, Cuevas O, Cercenado E, Marcos C, Bautista V, Castellares C, et al. Methicillin-resistant *Staphylococcus aureus* in Spain: Molecular epidemiology and utility of different typing methods. J Clin Microbiol. 2009;47:1620–7.

101. Albrecht N, Jatzwauk L, Slickers P, Ehricht R, Monecke S. Clonal replacement of epidemic methicillin-resistant *Staphylococcus aureus* strains in a German University Hospital over a period of eleven years. PLoS One. 2011;6:e28189.

102. Holden, M TG. Hsu, LY. Kurt, K. Weinert, LA. Mather, AE. Harris S, Strommenger, B. Layer F, Witte, W. Lencastre, HD. Skov R, Westh, H. Edgeworth, J. Gould I, Gant, V. Cooke, J. Edwards G, Mcadam, PR. Templeton K, et al. A genomic portrait of the emergence, evolution, and global spread of a methicillin-resistant *Staphylococcus aureus* pandemic. Genome Res. 2013;23:653–64.

103. Griggs DJ, Marona H, Piddock LJ V. Selection of moxifloxacin-resistant *Staphylococcus aureus* compared with five other fluoroquinolones. J Antimicrob Chemother. 2003;51:1403–1307.

104. Johnson AP, Aucken HM, Cavendish S, Ganner M, Wale MCJ, Warner M, et al. Dominance of EMRSA-15 and -16 among MRSA causing nosocomial bacteraemia in the UK: analysis of isolates from the European Antimicrobial Resistance Surveillance System (EARSS). J Antimicrob Chemother. 2001;48:143–4.

105. Kutter E, Sulakvelidze A. Bacteriophages : Biology and Applications. 2004.

106. Orlova EV. Bacteriophages and their structural organisation. Bacteriophages [Internet]. 2012;3–30. Available from:

http://www.intechopen.com/books/bacteriophages/bacteriophages-and-their-structural-organisation-

107. Oliveira H, Sampaio M, Melo LDR, Dias O, Pope WH, Hatfull GF, et al. Staphylococci phages display vast genomic diversity and evolutionary relationships. BMC Genomics. 2019;20:1–14.

108. Ackermann H-W. Bacteriophage observations and evolution. Res Microbiol. 2003;154:245–51.

109. Nelson D. Phage taxonomy: We agree to disagree. J Bacteriol. 2004;186:7029–31.

110. Lawrence JG, Hatfull GF, Hendrix RW. Imbroglios of viral taxonomy: Genetic exchange and failings of phenetic approaches. J Bacteriol. 2002;184.

111. Hendrix RW, Lawrence JG, Hatfull GF, Casjens S. The origins and ongoing evolution of viruses. Trends Microbiol. 2000;8:504–8.

112. Lavigne R, Darius P, Summer EJ, Seto D, Mahadevan P, Nilsson AS, et al. Classification of *Myoviridae* bacteriophages using protein sequence similarity. BMC Microbiol. 2009;9:1–16.

113. Walker PJ, Siddell SG, Lefkowitz EJ, Mushegian AR, Dempsey DM, Dutilh BE, et al. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2018). Arch Virol [Internet]. Springer Vienna; 2019;1–13. Available from: https://doi.org/10.1007/s00705-019-04306-w

114. 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. Syst Biol. 2019;

115. Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P. Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: Steps towards a unified taxonomy. J Gen Virol. 2018;99:1331–43.

116. Baess I. Report on a pseudolysogenic mycobacterium and a review of the literature concerning pseudolysogeny. Acta Pathol Microbiol Scand Sect B Microbiol

Immunol. 1971;79B:428–34.

117. Kazi M, Annapure US. Bacteriophage biocontrol of foodborne pathogens. J Food Sci Technol. 2016;53:1355–62.

118. Canchaya C, Proux C, Fournous G, Bruttin A, Brussow H. Prophage genomics. Microbiol Mol Biol Rev [Internet]. 2003;67:238–76. Available from: http://mmbr.asm.org/cgi/doi/10.1128/MMBR.67.2.238-276.2003

119. Atsumi S, Little JW. Role of the lytic repressor in prophage induction of phage as analyzed by a module-replacement approach. Proc Natl Acad Sci U S A [Internet]. 2006;103:4558–63. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1450210&tool=pmcentrez&rendertype=abstract

120. Abedon ST, Kuhl SJ, Blasdel BG, Kutter EM. Phage treatment of human infections. Bacteriophage. 2011;1:66–85.

121. Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML, Brüssow H. Phage as agents of lateral gene transfer. Curr Opin Microbiol. 2003;6:417–24.

122. Saunders JR, Allison H, James CE, McCarthy AJ, Sharp R. Phage-mediated transfer of virulence genes. J Chem Technol Biotechnol. 2001;76:662–6.

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

124. Coleman DC, Sullivan DJ, Russel RJ, Arbuthnott JP, Carey BF, Pomeroy HM. *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of -lysin, staphylokinase and enterotoxin a: molecular mechanism of triple conversion. Microbiology. 2009;135:1679–97.

125. Yuan Y, Gao M. Jumbo bacteriophages: An overview. Front Microbiol. 2017;8.

126. Klumpp J, Fouts DE, Sozhamannan S. Next generation sequencing technologies and the changing landscape of phage genomics. Bacteriophage. 2012;2:190–9.

127. Deghorain M, Van Melderen L. The staphylococci phages family: An overview. Viruses. 2012;4:3316–35.

128. Hatfull GF, Hendrix RW. Bacteriophages and their genomes. Curr Opin Virol. 2011;1:298–303.

129. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, et al. Origins of highly mosaic mycobacteriophage genomes. Cell. 2003;113:171–82.

130. Evolution P. Phage evolution. Bacteriophage Ecol. 2009. p. 77–194.

131. 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. Syst Biol. 2019;69:110–23.

132. Hendrix RW. Bacteriophage genomics. Curr Opin Microbiol. 2003;6:506–11.

133. Hatfull GF. Bacteriophage genomics. Curr Opin Microbiol. 2008;11:447–53.

134. Young R. Phage lysis: Three steps, three choices, one outcome. J Microbiol. 2014;52:243–58.

135. Hathaway H, Ajuebor J, Stephens L, Coffey A, Potter U, Sutton JM, et al. Thermally triggered release of the bacteriophage endolysin CHAPKand the bacteriocin lysostaphin for the control of methicillin resistant *Staphylococcus aureus* (MRSA). J Control Release [Internet]. The Authors; 2017;245:108–15. Available from: http://dx.doi.org/10.1016/j.jconrel.2016.11.030

136. Young R, Wang IN, Roof WD. Phages will out: Strategies of host cell lysis. Trends Microbiol. 2000;8:120–8.

137. Schmelcher M, Donovan DM, Loessner MJ. Bacteriophage endolysins as novel antimicrobials. Future Microbiol. 2012;7:1147–71.

138. Young R, Bläsi U. Holins: form and function in bacteriophage lysis. FEMS Microbiol Rev. 1995;17:191–205.

139. Grundling A, Manson MD, Young R. Holins kill without warning. Proc Natl Acad Sci. 2001;98:9348–52.

140. Wang I-N, Smith DL, Young R. Holins: The protein clocks of bacteriophage infections. Annu Rev Microbiol. 2000;54:799–825.

141. Fischetti VA. Bacteriophage endolysins: A novel anti-infective to control Gram-

positive pathogens. Int J Med Microbiol [Internet]. Elsevier GmbH.; 2010;300:357–62. Available from: http://dx.doi.org/10.1016/j.ijmm.2010.04.002

142. Cheng Q, Nelson D, Zhu S, Fischetti VA. Removal of group B streptococci colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. Antimicrob Agents Chemother. 2005;49:111–7.

143. Grandgirard D, Loeffler JM, Fischetti VA, Leib SL. Phage lytic enzyme cpl-1 for antibacterial therapy in experimental pneumococcal meningitis. J Infect Dis. 2008;197:1519–22.

144. Jado I, López R, García E, Fenoll A, Casal J, García P, et al. Phage lytic enzymes as therapy for antibiotic-resistant *Streptococcus pneumoniae* infection in a murine sepsis model. J Antimicrob Chemother. 2003;52:967–73.

145. Loeffler JM, Djurkovic S, Fischetti VA. Phage Lytic Enzyme Cpl-1 as a Novel Antimicrobial for Pneumococcal Bacteremia. Infect Immun. 2003;71:6199–204.

146. Horgan M, O'Flynn G, Garry J, Cooney J, Coffey A, Fitzgerald GF, et al. Phage lysin LysK can be truncated to its CHAP domain and retain lytic activity against live antibiotic-resistant staphylococci. Appl Environ Microbiol. 2009;75:872–4.

147. O'Flaherty S, Coffey A, Meaney W, Fitzgerald GF, Ross RP. The recombinant phage lysin LysK has a broad spectrum of lytic activity against clinically relevant staphylococci, including methicillin-resistant *Staphylococcus aureus*. J Bacteriol. 2005;187:7161–4.

148. Schmelcher M, Shen Y, Nelson DC, Eugster MR, Eichenseher F, Hanke DC, et al. Evolutionarily distinct bacteriophage endolysins featuring conserved peptidoglycan cleavage sites protect mice from MRSA infection. J Antimicrob Chemother. 2014;70:1453–65.

149. Piuri M, Hatfull GF. A peptidoglycan hydrolase motif within the mycobacteriophage TM4 tape measure protein promotes efficient infection of stationary phase cells. Mol Microbiol. 2006;62:1569–85.

150. Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B. Bacteriophages and Phage-Derived Proteins – Application Approaches. Curr Med Chem. 2015;22:1757–73.

151. Rodríguez L, Martínez B, Zhou Y, Rodríguez A, Donovan DM, García P. Lytic activity

of the virion-associated peptidoglycan hydrolase HydH5 of Staphylococcus aureus bacteriophage vB-SauS-phiIPLA88. BMC Microbiol. 2011;11.

152. Rodríguez-Rubio L, Martínez B, Donovan DM, Rodríguez A, García P. Bacteriophage virion-associated peptidoglycan hydrolases: Potential new enzybiotics. Crit Rev Microbiol. 2013;39:427–34.

153. Fischetti VA. Bacteriophage lytic enzymes: Novel anti-infectives. Trends Microbiol. 2005;13:491–6.

154. Pires DP, Oliveira H, Melo LDR, Sillankorva S, Azeredo J. Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl. Microbiol. Biotechnol. 2016.

155. Cornelissen A, Ceyssens PJ, T'Syen J, van Praet H, Noben JP, Shaburova O V., et al. The t7-related pseudomonas putida phage φ15 displays virion-associated biofilm degradation properties. PLoS One. 2011;6.

156. Knecht LE, Veljkovic M, Fieseler L. Diversity and Function of Phage Encoded Depolymerases. Front Microbiol. 2020;10.

157. Cornelissen A, Ceyssens PJ, Krylov VN, Noben JP, Volckaert G, Lavigne R. Identification of EPS-degrading activity within the tail spikes of the novel *Pseudomonas putida* phage AF. Virology [Internet]. Elsevier; 2012;434:251–6. Available from: http://dx.doi.org/10.1016/j.virol.2012.09.030

158. Cui Z, Guo X, Dong K, Zhang Y, Li Q, Zhu Y, et al. Safety assessment of *Staphylococcus* phages of the family *Myoviridae* based on complete genome sequences. Sci Rep. 2017;7:1–8.

159. Skurnik M, Pajunen M, Kiljunen S. Biotechnological challenges of phage therapy. Biotechnol Lett. 2007;29:995–1003.

160. Kim K-H, Chang H-W, Nam Y-D, Roh SW, Bae J-W. Phenotypic characterization and genomic analysis of the *Shigella sonnei* bacteriophage SP18. J Microbiol [Internet]. 2010;48:213–22. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20437154

161. Waldor MK, Mekalanos JJ. Lysogenic conversion by a filamentous phage encoding cholera toxin. Science (80-). 1996;272:1910–4.

162. Kutter E, De Vos D, Gvasalia G, Alavidze Z, Gogokhia L, Kuhl S, et al. Phage therapy in clinical practice: treatment of human infections. Curr Pharm Biotechnol. 2010;11:69–86.

163. Summers WC. Bacteriophage therapy. Annu Rev Microbiol. 2001;437–51.

164. Furfaro LL, Payne MS, Chang BJ. Bacteriophage therapy: Clinical trials and regulatory hurdles. Front Cell Infect Microbiol. 2018;8.

165. O'Flaherty S, Ross RP, Coffey A. Bacteriophage and their lysins for elimination of infectious bacteria: Review article. FEMS Microbiol Rev. 2009;33:801–19.

166. 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;5:226–35.

167. Barrow PA, Soothill JS. Bacteriophage therapy and prophylaxis: Rediscovery and renewed assessment of potential. Trends Microbiol. 1997;5:268–71.

168. Harper D, Parracho H, Walker J, Sharp R, Hughes G, Werthén M, et al. Bacteriophages and biofilms. Antibiotics [Internet]. 2014;3:270–84. Available from: http://www.mdpi.com/2079-6382/3/3/270/

169. Sutherland IW, Hughes KA, Skillman LC, Tait K. The interaction of phage and biofilms. FEMS Microbiol Lett. 2004;232:1–6.

170. McCallin S, Brüssow H. Phage therapy: an alternative or adjunct to antibiotics? Emerg Top Life Sci. 2017;1:105–16.

171. Chan BK, Abedon ST, Loc-Carrillo C. Phage cocktails and the future of phage therapy. Future Microbiol [Internet]. 2013;8:769–83. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23701332

172. Sulakvelidze A, Alavidze Z, Glenn Morris Jr. J, Summers WC. Bacteriophage therapy. Annu Rev Microbiol. 2001;437–51.

173. Fauconnier A. Regulating phage therapy. EMBO Rep. 2017;18:198–200.

174. Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. Nat Rev Microbiol [Internet]. Nature Publishing Group; 2010;8:317–27. Available from:

http://dx.doi.org/10.1038/nrmicro2315

175. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, Bickle TA, Bitinaite J, et al. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res. 2003;31:1805–12.

176. Marchfelder A. Special focus CRISPR-Cas. RNA Biol. 2013;10:655–8.

177. Leon LM, Mendoza SD, Bondy-Denomy J. How bacteria control the CRISPR-Cas arsenal. Curr Opin Microbiol. 2018;42:87–95.

178. Barrangou R, Marraffini LA. CRISPR-cas systems: Prokaryotes upgrade to adaptive immunity. Mol Cell. 2014;54:234–44.

179. Mojica FJM, Díez-Villaseñor C, Soria E, Juez G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. Mol Microbiol. 2000;36:244–6.

180. Trasanidou D, Gerós AS, Mohanraju P, Nieuwenweg AC, Nobrega FL, Staals RHJ. Keeping crispr in check: diverse mechanisms of phage-encoded anti-crisprs. FEMS Microbiol Lett. 2019;366.

181. Bertozzi Silva J, Storms Z, Sauvageau D. Host receptors for bacteriophage adsorption. FEMS Microbiol Lett. 2016;363:1–11.

182. Seed KD. Battling phages: How bacteria defend against viral attack. PLoS Pathog. 2015;11:1–5.

183. Dy RL, Richter C, Salmond GPC, Fineran PC. Remarkable mechanisms in microbes to resist phage infections. Annu Rev Virol. 2014;1:307–31.

184. Orzechowska B, Mohammed M. The war between bacteria and bacteriophages. Grow Handl Bact Cult. 2019;

185. Moller AG, Lindsay JA, Read TD. Determinants of phage host range in Staphylococcus species. Appl Environ Microbiol. 2019;85:1–16.

186. Ram G, Chen J, Ross HF, Novick RP, Musser JM. Precisely modulated pathogenicity island interference with late phage gene transcription. Proc Natl Acad Sci U S A. 2014;111:14536–41.

187. Alves DR, Gaudion A, Bean JE, Perez Esteban P, Arnot TC, Harper DR, et al. Combined use of bacteriophage K and a novel bacteriophage to reduce *Staphylococcus aureus*biofilm formation. Appl Environ Microbiol. 2014;80:6694–703.

188. Gutiérrez D, Vandenheuvel D, Martínez B, Rodríguez A, Lavigne R, García P. Two phages, phiIPLA-RODI and phiIPLA-C1C, lyse mono-and dual-species staphylococcal biofilms. Appl Environ Microbiol. 2015;81:3336–48.

189. Kelly D, Mcauliffe O, Ross RP, Coffey A. Prevention of *Staphylococcus aureus*biofilm formation and reduction in established biofilm density using a combination of phage K and modified derivatives. Lett Appl Microbiol. 2012;54:286–91.

190. Melo LDR, Brandão A, Akturk E, Santos SB, Azeredo J. Characterization of a new *Staphylococcus aureus Kayvirus* harboring a lysin active against biofilms. Viruses. 2018;10:182.

191. Mapes AC, Trautner BW, Liao KS, Ramig RF. Development of expanded host range phage active on biofilms of multi-drug resistant *Pseudomonas aeruginosa*. Bacteriophage. Taylor & Francis; 2016;6:e1096995.

192. Otto M. Staphylococcal biofilms. Curr Top Microbiol Immunol. 2008;322:207–28.

193. Hansen MF, Svenningsen S Lo, Røder HL, Middelboe M, Burmølle M. Big Impact of the Tiny: Bacteriophage–Bacteria Interactions in Biofilms. Trends Microbiol. 2019;27:739–52.

194. Suresh MK, Biswas R, Biswas L. An update on recent developments in the prevention and treatment of *Staphylococcus aureus* biofilms. Int J Med Microbiol. 2019;309:1–12.

195. Flemming HC, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. Biofilms: An emergent form of bacterial life. Nat Rev Microbiol. 2016;14:563–75.

196. Harrison JJ, Turner RJ, Ceri H. Persister cells, the biofilm matrix and tolerance to metal cations in biofilm and planktonic *Pseudomonas aeruginosa*. Environ Microbiol. 2005;7:981–94.

197. Lu TK, Collins JJ. Dispersing biofilms with engineered enzymatic bacteriophage.

Proc Natl Acad Sci [Internet]. 2007;104:11197–202. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0704624104

198. Gutiérrez D, Ruas-Madiedo P, Martínez B, Rodríguez A, García P. Effective removal of staphylococcal biofilms by the endolysin LysH5. PLoS One. 2014;9:e107307.

199. Donlan RM, Costerton JW. Biofilms: survival mechanisms of clinically relevant microorganisms. Clin Microbiol Rev. 2002;15:167–93.

200. Hood AM. Phage typing of *Staphylococcus aureus*. J Hyg (Lond). 1953;51:1–15.

201. Kutateladze M, Adamia R. Phage therapy experience at the Eliava Institute. Med Mal Infect. 2008;38:426–30.

202. Łobocka M, Hejnowicz MS, Dabrowski K, Gozdek A, Kosakowski J, Witkowska M, et al. Genomics of Staphylococcal Twort-like Phages - Potential Therapeutics of the Post-Antibiotic Era. Adv Virus Res. 2012;83:143–216.

203. O'Flaherty S, Ross RP, Meaney W, Fitzgerald GF, Elbreki MF, Coffey A. Potential of the polyvalent anti-Staphylococcus bacteriophage K for control of antibiotic-resistant staphylococci from hospitals. Appl Environ Microbiol. 2005;71:1836–42.

204. Hsieh SE, Lo HH, Chen ST, Lee MC, Tseng YH. Wide host range and strong lytic activity of *Staphylococcus aureus* lytic phage Stau2. Appl Environ Microbiol. 2011;77:756–61.

205. Kraushaar B, Thanh MD, Hammerl JA, Reetz J, Fetsch A, Hertwig S. Isolation and characterization of phages with lytic activity against methicillin-resistant *Staphylococcus aureus* strains belonging to clonal complex 398. Arch Virol. 2013;158:2341–50.

206. Yu P, Mathieu J, Li M, Dai Z, Alvarez PJJ. Isolation of polyvalent bacteriophages by sequential multiple-host approaches. Appl Environ Microbiol. 2015;82:808–15.

207. Azam AH, Tanji Y. Peculiarities of *Staphylococcus aureus* phages and their possible application in phage therapy. Appl Microbiol Biotechnol. 2019;103:4279–89.

208. Xia G, Wolz C. Phages of *Staphylococcus aureus* and their impact on host evolution. Infect Genet Evol [Internet]. Elsevier B.V.; 2014;21:593–601. Available from: http://dx.doi.org/10.1016/j.meegid.2013.04.022

209. Xia G, Maier L, Sanchez-Carballo P, Li M, Otto M, Holst O, et al. Glycosylation of wall teichoic acid in *Staphylococcus aureus* by TarM. J Biol Chem. 2010;285:13405–15.

210. Xia G, Corrigan RM, Winstel V, Goerke C, Gründling A, Peschel A. Wall teichoic acid-dependent adsorption of staphylococcal siphovirus and myovirus. J Bacteriol. 2011;193:4006–9.

211. Vandersteegen K, Kropinski AM, Nash JHE, Noben J-P, Hermans K, Lavigne R. Romulus and Remus, two phage isolates representing a distinct clade within the *twortlikevirus* genus, display suitable properties for phage therapy. J Virol [Internet]. 2013;87:3237–47. Available from: http://jvi.asm.org/cgi/doi/10.1128/JVI.02763-12

212. Drilling A, Morales S, Jardeleza C, Vreugde S, Speck P, Wormald PJ. Bacteriophage reduces biofilm of *Staphylococcus aureus ex vivo* isolates from chronic rhinosinusitis patients. Am J Rhinol Allergy. 2014;28:3–11.

213. Geredew Kifelew L, Mitchell JG, Speck P. Mini-review: efficacy of lytic bacteriophages on multispecies biofilms. Biofouling. 2019;39:472–81.

214. Knezevic P, Curcin S, Aleksic V, Petrusic M, Vlaski L. Phage-antibiotic synergism: A possible approach to combatting *Pseudomonas aeruginosa*. Res Microbiol [Internet]. Elsevier Masson SAS; 2013;164:55–60. Available from: http://dx.doi.org/10.1016/j.resmic.2012.08.008

215. Ryan EM, Alkawareek MY, Donnelly RF, Gilmore BF. Synergistic phage-antibiotic combinations for the control of *Escherichia coli* biofilms in vitro. FEMS Immunol Med Microbiol. 2012;65:395–8.

216. Daniel A, Euler C, Collin M, Chahales P, Gorelick KJ, Fischetti VA. Synergism between a novel chimeric lysin and oxacillin protects against infection by methicillin-resistant *Staphylococcus aureus*. Antimicrob Agents Chemother. 2010;54:1603–12.

217. Rahman M, Kim SSM, Kim SSM, Seol SY, Kim J. Characterization of induced *Staphylococcus aureus* bacteriophage SAP-26 and its anti-biofilm activity with rifampicin. Biofouling. 2011;27:1087–93.

218. Kumaran D, Taha M, Yi QL, Ramirez-Arcos S, Diallo JS, Carli A, et al. Does treatment order matter? Investigating the ability of bacteriophage to augment antibiotic activity

against *Staphylococcus aureus* biofilms. Front Microbiol. 2018;9.

219. Tkhilaishvili T, Lombardi L, Klatt AB, Trampuz A, Di Luca M. Bacteriophage Sb-1 enhances antibiotic activity against biofilm, degrades exopolysaccharide matrix and targets persisters of *Staphylococcus aureus*. Int J Antimicrob Agents. 2018;52:842–53.

220. Klumpp J, Lavigne R, Loessner MJ, Ackermann H-W. The SPO1-related bacteriophages. Arch Virol. 2010;155:1547–61.

221. Kropinski AM, Mazzocco A, Waddell TE, Johnson RP. Enumeration of bacteriophages by double agar overlay plaque assay. Bacteriophages Methods Protoc Vol 1 Isol Charact Interact [Internet]. 2009. p. 69–76. Available from: http://link.springer.com/10.1007/978-1-60327-565-1

222. Raya RR, H'bert EM. Isolation of phage via induction of lysogens. Bacteriophages Biol Appl. Humana Press; 2009. p. 23–32.

223. Stepanovic S, Vukovic D, Dakic I, Savic B, Svabic-Vlahovic M. A modified microtiter - plate test for quantification of staphylococcal biofilm formation. J Microbiol Methods [Internet]. 2000;40:175–9. Available from: https://www.researchgate.net/profile/Branislava_Savic/publication/12615362_A_mo dified_microtiter- plate_test_for_quantification_of_staphylococcal_biofilm_formation/links/5666d8510 8ae4d38f7ac0067.pdf

224. Croes S, Deurenberg RH, Boumans MLL, Beisser PS, Neef C, Stobberingh EE. *Staphylococcus aureus* biofilm formation at the physiologic glucose concentration depends on the S. aureus lineage. BMC Microbiol. 2009;9:1–9.

225. Cafiso V, Bertuccio T, Santagati M, Demelio V, Spina D, Nicoletti G, et al. agr-Genotyping and transcriptional analysis of biofilm-producing *Staphylococcus aureus*. FEMS Immunol Med Microbiol. 2007;51:220–7.

226. Gutiérrez D, Fernández L, Martínez B, Ruas-Madiedo P, García P, Rodríguez A. Real-time assessment of *Staphylococcus aureus* biofilm disruption by phage-derived proteins. Front Microbiol. 2017;8:1–10.

227. Tasse J, Trouillet-Assant S, Josse J, Martins-Simões P, Valour F, Langlois-Jacques C,

et al. Association between biofilm formation phenotype and clonal lineage in *Staphylococcus aureus* strains from bone and joint infections. PLoS One. 2018;13:1–11.

228. Rihtman B, Meaden S, Clokie MRJ, Koskella B, Millard AD. Assessing Illumina technology for the high-throughput sequencing of bacteriophage genomes. PeerJ [Internet]. 2016;4:e2055. Available from: https://peerj.com/articles/2055

229. Andrews S. FastQC - A quality control tool for high throughput sequence data. Available at http://http://www.bioinformatics.babraham.ac.uk/projects/fastqc. 2010.

230. Joshi N, Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available at https://github.com/najoshi/sickle. 2011;2011.

231. Krueger F. Trim Galore! (Version 0.4.3) [Software]. Available at https://github.com/FelixKrueger/TrimGalore [Internet]. http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/; Available from: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

232. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

233. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.

234. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol [Internet]. 2012;19:455–77. Available from: http://online.liebertpub.com/doi/abs/10.1089/cmb.2012.0021

235. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: Interactive visualization of *de novo* genome assemblies. Bioinformatics. 2015;31:3350–2.

236. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: Sequence visualization and annotation. Bioinformatics. 2000;16:944–5.

237. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.

238. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

239. Hildebrand A, Remmert M, Biegert A, Soding J. Fast and accurate automatic structure prediction with HHpred. Proteins Struct Funct Bioinforma. 2009;77:128–32.

240. Finn RD, Clements J, Eddy SR. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Res. 2011;39:W29–37.

241. Katz L, Griswold T, Morrison S, Caravas J, Zhang S, Bakker H, et al. Mashtree: a rapid comparison of whole genome sequence files. J Open Source Softw. 2019;4:1762.

242. Rambaut A. FigTree v1.4.3 [Internet]. 2009. Available from: http://tree.bio.ed.ac.uk/software/figtree/

243. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol [Internet]. Genome Biology; 2016;17:1–14. Available from: http://dx.doi.org/10.1186/s13059-016-0997-x

244. Galili T, O'Callaghan A, Sidi J, Sievert C, O'Callaghan A, Sidi J, et al. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. Bioinformatics [Internet]. 2017;34:1600–2. Available from: http://academic.oup.com/bioinformatics/advance-article/doi/10.1093/bioinformatics/btx657/4562328

245. Rstudio Team. RStudio: Integrated development for R. [Online] RStudio, Inc., Boston, MA. 2016.

246. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31:3691–3.

247. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. Phandango: An interactive viewer for bacterial population genomics. Bioinformatics. 2018;34:292–3.

248. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics. 2011;12.

249. Sullivan MJ, Petty NK, Beatson SA. Easyfig: A genome comparison visualizer.

Bioinformatics. 2011;27:1009–10.

250. Steinegger M, Meier M, Mirdita M, Voehringer H, Haunsberger SJ, Soeding J. HH-suite3 for fast remote homology detection and deep protein annotation. bioRxiv. 2019;20.

251. Rosenstein R, Götz F. Genomic differences between the food-grade *Staphylococcus carnosus* and pathogenic staphylococcal species. Int J Med Microbiol. 2010;300:104–8.

252. Finch RG, Eliopoulos GM. Safety and efficacy of glycopeptide antibiotics. J Antimicrob Chemother. 2005;55:ii5–113.

253. Welte T, Pletz MW. Antimicrobial treatment of nosocomial meticillin-resistant *Staphylococcus aureus* (MRSA) pneumonia: Current and future options. Int J Antimicrob Agents [Internet]. Elsevier B.V.; 2010;36:391–400. Available from: http://dx.doi.org/10.1016/j.ijantimicag.2010.06.045

254. Sievert DM, Rudrik JT, Patel JB, McDonald LC, Wilkins MJ, Hageman JC. Vancomycin-Resistant *Staphylococcus aureus* in the United States, 2002-2006. Clin Infect Dis [Internet]. 2008;46:668–74. Available from: https://academic.oup.com/cid/article-lookup/doi/10.1086/527392

255. Chang S, Sievert DM, Hageman JC, Boulton ML, Tenover FC, Downes FP, et al. Infection with vancomycin-resistant *Staphylococcus aureus* containing the vanA resistance gene. N Engl J Med. 2003;348:1342–7.

256. Hiramatsu K, Hanaki H, Ino T, Yabuta K, Oguri T, Tenover FC. Methicillin-resistant *Staphylococcus aureus* clinical strain with reduced vancomycin susceptibility. J Antimicrob Chemother. 1997;40:135–46.

257. Foster TJ. Antibiotic resistance in *Staphylococcus aureus*. Current status and future prospects. FEMS Microbiol Rev. 2017;41:430–49.

258. Choo EJ, Chambers HF. Treatment of methicillin-resistant *Staphylococcus aureus* bacteremia. Infect Chemother. 2016;48:267–73.

259. Humphries RM, Pollett S, Sakoulas G. A current perspective on daptomycin for the clinical microbiologist. Clin Microbiol Rev. 2013;26:759–80.

260. Czaplewski L, Bax R, Clokie M, Dawson M, Fairhead H, Fischetti VA, et al. Alternatives to antibiotics-a pipeline portfolio review. Lancet Infect Dis. 2016;16:239–51.

261. Weber-Dabrowska B, Jończyk-Matysiak E, Zaczek M, Łobocka M, Łusiak-Szelachowska M, Górski A. Bacteriophage procurement for therapeutic purposes. Front Microbiol. 2016;7:1–14.

262. Hyman P, Abedon ST. Practical methods for determining phage growth parameters. Bacteriophages Methods Protoc Vol 1 Isol Charact Interact [Internet]. 2009. p. 175–202. Available from: http://link.springer.com/10.1007/978-1-60327-565-1

263. Gill J, Young RF. Therapeutic applications of phage biology: history, practice and recommendations. Emerg Trends Antibact Discov Answering Call to Arms (Edited by Alita A Mill Paul F Miller). Caiter Academic Press, UK; 2011.

264. Gill JJ, Hyman P. Phage choice isolation and preperation for phage therapy. Curr Pharm Biotechnol. 2010;11:2–14.

265. Clokie MRJ, Kropinski AM. Bacteriophages Methods and Protocols Volume 1: Isolation, Characterization, and Interactions [Internet]. Life Sci. 2009. Available from: http://books.google.com/books?id=Ku2wPAAACAAJ

266. Loc-Carrillo C, Abedon ST. Pros and cons of phage therapy. Bacteriophage [Internet]. 2011;1:111–4. Available from: http://www.tandfonline.com/doi/abs/10.4161/bact.1.2.14590

267. Mattila S, Ruotsalainen P, Jalasvuori M. On-demand isolation of bacteriophages against drug-resistant bacteria for personalized phage therapy. Front Microbiol. 2015;6:1–7.

268. Gupta R, Prasad Y. Efficacy of polyvalent bacteriophage P-27/HP to control multidrug resistant *Staphylococcus aureus* associated with human infections. Curr Microbiol. 2011;62:255–60.

269. Wang Z, Zheng P, Ji W, Fu Q, Wang H, Yan Y, et al. SLPW: A virulent bacteriophage targeting methicillin-resistant *staphylococcus aureus in vitro* and *in vivo*. Front

Microbiol. 2016;7.

270. El Haddad L, Roy JP, Khalil GE, St-Gelais D, Champagne CP, Labrie S, et al. Efficacy of two *Staphylococcus aureus* phage cocktails in cheese production. Int J Food Microbiol [Internet]. Elsevier B.V.; 2016;217:7–13. Available from: http://dx.doi.org/10.1016/j.ijfoodmicro.2015.10.001

271. Fish R, Kutter E, Wheat G, Blasdel B, Kutateladze M, Kuhl S. Bacteriophage treatment of intransigent diabetic toe ulcers: a case series. J Wound Care [Internet]. 2016;25:S27–33. Available from: http://www.magonlinelibrary.com/doi/10.12968/jowc.2016.25.7.S27

272. Leszczyński P, Weber-Dąbrowska B, Kohutnicka M, Łuczak M, Górecki A, Górski A. Successful eradication of methicillin-resistant *Staphylococcus aureus* (MRSA) intestinal carrier status in a healthcare worker - Case report. Folia Microbiol (Praha). 2006;51:236–8.

273. Verstappen KM, Tulinski P, Duim B, Fluit AC, Carney J, Van Nes A, et al. The effectiveness of bacteriophages against methicillin-resistant *Staphylococcus aureus* ST398 nasal colonization in pigs. PLoS One. 2016;11:1–10.

274. Estrella LA, Quinones J, Henry M, Hannah RM, Pope RK, Hamilton T, et al. Characterization of novel *Staphylococcus aureus* lytic phage and defining their combinatorial virulence using the OmniLog® system. Bacteriophage [Internet]. 2016;6:e1219440. Available from: https://www.tandfonline.com/doi/full/10.1080/21597081.2016.1219440

275. Adriaenssens EM, Rodney Brister J. How to name and classify your phage: An informal guide. Viruses. 2017;9:1–9.

276. Ly-Chatain MH. The factors affecting effectiveness of treatment in phages therapy. Front Microbiol. 2014;5:1–7.

277. Mirzaei MK, Nilsson AS. Isolation of phages for phage therapy: A comparison of spot tests and efficiency of plating analyses for determination of host range and efficacy. PLoS One. 2015;10:1–13.

278. Abedon ST. Lysis from without. Bacteriophage [Internet]. 2011;1:46–9. Available

from: http://www.tandfonline.com/doi/abs/10.4161/bact.1.1.13980

279. Brüssow H, Kutter E. Phage ecology. Bacteriophages Biol Appl. 2004. p. 129–64.

280. Lee YD, Park JH. Characterization and application of phages isolated from sewage for reduction of *Escherichia coli* O157: H7 in biofilm. LWT - Food Sci Technol [Internet]. Elsevier Ltd; 2015;60:571–7. Available from: http://dx.doi.org/10.1016/j.lwt.2014.09.017

281. Ribeiro KVG, Ribeiro C, Dias RS, Cardoso SA, de Paula SO, Zanuncio JC, et al. Bacteriophage isolated from sewage eliminates and prevents the establishment of *Escherichia coli* biofilm. Adv Pharm Bull. 2018;8:85–95.

282. Sangha KK, Kumar BVS, Agrawal RK, Deka D, Verma R. Proteomic characterization of lytic bacteriophages of *Staphylococcus aureus* isolated from sewage affluent of India. Int Sch Res Not. 2014;2014:1–7.

283. Synnott AJ, Kuang Y, Kurimoto M, Yamamichi K, Iwano H, Tanji Y. Isolation from sewage influent and characterization of novel *Staphylococcus aureus* bacteriophages with wide host ranges and potent lytic capabilities. Appl Environ Microbiol. 2009;75:4483–90.

284. Rosenstein R, Nerz C, Biswas L, Resch A, Raddatz G, Schuster SC, et al. Genome analysis of the meat starter culture bacterium *Staphylococcus carnosus* TM300. Appl Environ Microbiol. 2009;75:811–22.

285. El Haddad L, Abdallah N Ben, Plante PL, Dumaresq J, Katsarava R, Labrie S, et al. Improving the safety of *Staphylococcus aureus* polyvalent phages by their production on a *Staphylococcus xylosus*strain. PLoS One. 2014;9:1–10.

286. Gutiérrez D, Martínez B, Rodríguez A, García P. Isolation and characterization of bacteriophages infecting *Staphylococcus epidermis*. Curr Microbiol. 2010;61:601–8.

287. Gutiérrez D, Briers Y, Rodríguez-Rubio L, Martínez B, Rodríguez A, Lavigne R, et al. Role of the pre-neck appendage protein (Dpo7) from phage vB_SepiS-phiIPLA7 as an anti-biofilm agent in staphylococcal species. Front Microbiol. 2015;6:1–10.

288. Chhibber S, Bansal S, Kaur S. Disrupting the mixed-species biofilm of *Klebsiella pneumoniae* B5055 and *Pseudomonas aeruginosa* PAO using bacteriophages alone or

in combination with xylitol. Microbiol (United Kingdom). 2015;161:1369–77.

289. Kutter E. Phage host range and efficiency of plating. Bacteriophages Methods Protoc Vol 1 Isol Charact Interact [Internet]. 2009. p. 141–9. Available from: http://link.springer.com/10.1007/978-1-60327-565-1

290. Jensen EC, Schrader HS, Rieland B, Thomas L, Lee KW, Nickerson KW, et al. Prevalence of Broad-Host-Range Lytic Bacteriophages of *Sphaerotilus natans* , *Escherichia coli*, and *Pseudomonas aeruginosa*. Appl Environ Microbiol. 1998;64:575–80.

291. Hamdi S, Rousseau GM, Labrie SJ, Tremblay DM, Kourda RS, Slama K Ben, et al. Characterization of two polyvalent phages infecting *Enterobacteriaceae*. Sci Rep [Internet]. Nature Publishing Group; 2017;7:1–12. Available from: http://dx.doi.org/10.1038/srep40349

292. Yu P, Mathieu J, Lu GW, Gabiatti N, Alvarez PJ. Control of antibiotic-resistant bacteria in activated sludge using polyvalent phages in conjunction with a production host. Environ Sci Technol Lett. 2017;4:137–42.

293. Pantůček R, Rosypalová A, Doškař J, Kailerová J, Růžičková V, Borecká P, et al. The polyvalent staphylococcal phage phi812: Its host-range mutants and related phages. Virology. 2008;246:241–52.

294. Vandersteegen K, Mattheus W, Ceyssens PJ, Bilocq F, de Vos D, Pirnay JP, et al. Microbiological and molecular assessment of bacteriophage ISP for the control of *Staphylococcus aureus*. PLoS One. 2011;6:e24418.

295. Bigwood T, Hudson JA, Billington C. Influence of host and bacteriophage concentrations on the inactivation of food-borne pathogenic bacteria by two phages: Research Letter. FEMS Microbiol Lett. 2009;291:59–64.

296. Abedon ST. Lysis of lysis-inhibited bacteriophage T4-infected cells. J Bacteriol. 1992;174:8073–80.

297. Beeton ML, Alves DR, Enright MC, Jenkins ATA. Assessing phage therapy against *Pseudomonas aeruginosa* using a *Galleria mellonella* infection model. Int J Antimicrob Agents [Internet]. Elsevier B.V.; 2015;46:196–200. Available from:

http://dx.doi.org/10.1016/j.ijantimicag.2015.04.005

298. Cui Z, Feng T, Gu F, Li Q, Dong K, Zhang Y, et al. Characterization and complete genome of the virulent *Myoviridae* phage JD007 active against a variety of S*taphylococcus aureus* isolates from different hospitals in Shanghai, China. Virol J. Virology Journal; 2017;14:1–8.

299. Middelboe M. Bacterial growth rate and marine virus-host dynamics. Microb Ecol [Internet]. 2000;40:114–24. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11029080

300. Avrani S, Lindell D. Convergent evolution toward an improved growth rate and a reduced resistance range in *Prochlorococcus* strains resistant to phage. Proc Natl Acad Sci. 2015;112:E2191–200.

301. Hall AR, Scanlan PD, Buckling A. Bacteria-phage coevolution and the emergence of generalist pathogens. Am Nat [Internet]. 2011;177:44–53. Available from: http://www.journals.uchicago.edu/doi/10.1086/657441

302. Lenski R. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. Int J Org Evol. 1988;42:425–32.

303. Reuter S, Török ME, Holden MTG, Reynolds R, Raven KE, Blane B, et al. Building a genomic framework for prospective MRSA surveillance in the United Kingdom and the Republic of Ireland Running Title : MRSA surveillance in the UK and Ireland Keywords : MRSA , surveillance , genomics. Genome Med. 2015;1–8.

304. Monecke S, Coombs G, Shore AC, Coleman DC, Akpaka P, Borg M, et al. A field guide to pandemic, epidemic and sporadic clones of methicillin-resistant *Staphylococcus aureus*. PLoS One. 2011;6:e17936.

305. Endersen L, Buttimer C, Nevin E, Coffey A, Neve H, Oliveira H, et al. Investigating the biocontrol and anti-biofilm potential of a three phage cocktail against *Cronobacter sakazakii* in different brands of infant formula. Int J Food Microbiol. 2017;253:1–11.

306. Hall-Stoodley L, Costerton JW, Stoodley P. Bacterial biofilms: from the natural environment to infectious diseases. Nat Rev Microbiol. 2004;2:95–108.

307. Lacroix-Gueu P, Briandet R, Lévêque-Fort S, Bellon-Fontaine MN, Fontaine-Aupart

MP. In situ measurements of viral particles diffusion inside mucoid biofilms. Comptes Rendus - Biol. 2005;328:1065–72.

308. Review M. Bacteriophages in biofilm control. EC. 2017;2:47–52.

309. Son JS, Lee SJ, Jun SY, Yoon SJ, Kang SH, Paik HR, et al. Antibacterial and biofilm removal activity of a podoviridae *Staphylococcus aureus* bacteriophage SAP-2 and a derived recombinant cell-wall-degrading enzyme. Appl Microbiol Biotechnol. 2010;86:1439–49.

310. Narasimhaiah MH, Asrani JY, Palaniswamy SM, Bhat J, George SE, Srinivasan R, et al. Therapeutic potential of staphylococcal bacteriophages for nasal decolonization of *Staphylococcus aureus* in mice. Adv Microbiol. 2013;2013:52–60.

311. Valle J, Toledo-arana A, Ghigo J-M, Amorena B, Penadés JR, Lasa I. SarA and not sigmaB is essential for biofilm development by *Staphylococcus aureus*. Mol Microbiol. 2003;48:1075–87.

312. Karaolis DKR, Rashid MH, Chythanya R, Luo W, Hyodo M, Hayakawa Y. c-di-GMP (3′-5′-cyclic diguanylic acid) inhibits *Stapylococcus aureus* cell-cell interactions and biofilm formation. Antimicrob Agents Chemother. 2005;49:1029–38.

313. Haney EF, Trimble MJ, Cheng JT, Vallé Q, Hancock REW. Critical assessment of methods to quantify biofilm growth and evaluate antibiofilm activity of host defence peptides. Biomolecules. 2018;8:29.

314. Reffuveille F, Josse J, Vallé Q, Mongaret C, Gangloff SC. *Staphylococcus aureus* biofilms and their impact on the medical field. Rise Virulence Antibiot Resist Staphylococcus aureus. 2017. p. 197–214.

315. Azeredo J, Azevedo NF, Briandet R, Cerca N, Coenye T, Costa AR, et al. Critical review on biofilm methods. Crit Rev Microbiol. 2017;43:313–51.

316. Foster TJ, Geoghegan JA, Ganesh VK, Hook M. Adhesion, invasion and evasion: The many functions of the surface proteins of *Staphylococcus aureus*. Nat Rev Microbiol. 2017;12:49–62.

317. Nourbakhsh F, Namvar AB. Detection of genes involved in biofilm formation in *Staphylococcus aureus* isolates. GMS Hyg Infect Control. 2016;11:1–5.

318. Fitzpatrick F, Humphrey H, O'Gara J. Environmental regulation of biofilm development in methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* clinical isolates. J Hosp Infect. 2006;62:120–2.

319. Fitzpatrick F, Humphreys H, Gara JPO. Evidence for icaADBC -independent biofilm development mechanism in methicillin-resistant *Staphylococcus aureus* clinical isolates. J Clin Microbiol. 2005;43:1973–6.

320. Cramton SE, Ulrich M, Go F. Anaerobic conditions induce expression of polysaccharide intercellular adhesin in *Staphylococcus aureus* and *Staphylococcus epidermis*. Infect Immun. 2001;69:4079–85.

321. Fluckiger U, Ulrich M, Steinhuber A, Do G, Mack D, Landmann R, et al. Biofilm formation, icaADBC transcription , and polysaccharide intercellular adhesin synthesis by staphylococci in a device-related infection model. Infect Immun. 2005;73:1811–9.

322. Seidl K, Goerke C, Wolz C, Mack D, Berger-ba B. *Staphylococcus aureus* CcpA affects biofilm formation. Infect Immun. 2008;76:2044–50.

323. Neill EO, Pozzi C, Houston P, Humphreys H, Robinson DA, Loughman A, et al. A novel *Staphylococcus aureus* biofilm phenotype mediated by the fibronectin-binding proteins , FnBPA and FnBPB. J Bacteriol. 2008;190:3835–50.

324. Luther MK, Parente DM, Caffrey AR, Daffinee KE, Lopes V V, Martin ET, et al. Clinical and genetic risk factors for biofilm-forming *Staphylococcus aureus*. Antimicrob Agents Chemother. 2018;62:1–10.

325. Dickey J, Perrot V. Adjunct phage treatment enhances the effectiveness of low antibiotic concentration against *Staphylococcus aureus* biofilms *in vitro*. PLoS One. 2019;14:1–17.

326. León M, Bastías R. Virulence reduction in bacteriophage resistant bacteria. Front Microbiol. 2015;6:1–7.

327. 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:e31698.

328. Lopes A, Pereira C, Almeida A. Sequential combined effect of phages and antibiotics on the inactivation of *Escherichia coli*. Microorganisms. 2018;6:125.

329. Örmälä A-M, Jalasvuori M. Phage therapy: Should bacterial resistance to phages be a concern, even in the long run? Bacteriophage. 2013;3:e24219.

330. Inal JM. Phage therapy: a reappraisal of bacteriophages as antibiotics. Arch Immunol Ther Exp (Warsz). 2003;51:237–44.

331. Mizoguchi K, Morita M, Fischer CR, Yoichi M, Tanji Y, Unno H. Coevolution of bacteriophage PP01 and *Escherichia coli* O157:H7 in continuous culture. Appl Environ Microbiol. 2003;66:170–6.

332. O'Flynn G, Coffey A, Fitzgerald G, Paul Ross R. *Salmonella enterica* phage-resistant mutant colonies display an unusual phenotype in the presence of phage Felix 01. Lett Appl Microbiol. 2007;33:801–19.

333. Kocharunchitt C, Ross T, McNeil DL. Use of bacteriophages as biocontrol agents to control Salmonella associated with seed sprouts. Int J Food Microbiol. 2009;128:453–9.

334. Mavrich TN, Hatfull GF. Bacteriophage evolution differs by host, lifestyle and genome. Nat Microbiol [Internet]. Nature Publishing Group; 2017;2:1–9. Available from: http://dx.doi.org/10.1038/nmicrobiol.2017.112

335. Barylski J, Dutilh BE, Schuller MBP, Edwards RA, Gillis A, Klumpp J, et al. Genomic, proteomic, and phylogenetic analysis of spounaviruses indicates paraphyly of the order Caudovirales. bioRxiv. 2017;1–44.

336. Veesler D, Cambillau C. A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. Microbiol Mol Biol Rev [Internet]. 2011;75:423–33. Available from: http://mmbr.asm.org/cgi/doi/10.1128/MMBR.00014-11

337. Casjens SR. Comparative genomics and evolution of the tailed-bacteriophages. Curr Opin Microbiol. 2005;8:451–8.

338. Kwan T, Liu J, DuBow M, Gros P, Pelletier J. The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophagess. Proc Natl Acad Sci [Internet]. 2005;102:5174–9. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.0501140102

339. Kwan T, Liu J, Dubow M, Gros P, Pelletier J. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. J Bacteriol [Internet]. 2006;188:1184–7. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1347338/%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1347338&tool=pmcentrez&rendertype=abstract

340. González-Candelas F, Comas I, Martínez JL, Galán JC, Baquero F. The evolution of antibiotic resistance. Genet Evol Infect Dis Second Ed. 2017;257–84.

341. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, et al. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife. 2015;4:e06416.

342. Ha AD, Denver DR. Comparative genomic analysis of 130 bacteriophages infecting bacteria in the genus *Pseudomonas*. Front Immunol. 2018;9:1–13.

343. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, et al. Comparative genomic analysis of 60 mycobacteriophage genomes: genome clustering, gene acquisition, and gene size. J Mol Biol [Internet]. Elsevier Ltd; 2010;397:119–43. Available from: http://dx.doi.org/10.1016/j.jmb.2010.01.011

344. Klumpp J, Fouts DE, Sozhamannan S. Bacteriophage functional genomics and its role in bacterial pathogen detection. Brief Funct Genomics. 2013;12:354–65.

345. Oliveira H, Melo LDR, Santos SB, Nobrega FL, Ferreira EC, Cerca N, et al. Molecular aspects and comparative genomics of bacteriophage endolysins. J Virol [Internet]. 2013;87:4558–70. Available from: http://jvi.asm.org/cgi/doi/10.1128/JVI.03277-12

346. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016;44:W16–21.

347. Leplae R, Lima-Mendez G, Toussaint A. ACLAME: A CLAssification of mobile genetic elements, update 2010. Nucleic Acids Res. 2009;

348. Lavigne R, Vandersteegen K. Group I introns in Staphylococcus bacteriophages. Future Virol. 2013;8:997–1005.

349. Cohen-Gonsaud M, Barthe P, Bagnéris C, Henderson B, Ward J, Roumestand C, et

al. The structure of a resuscitation-promoting factor domain from *Mycobacterium tuberculosis* shows homology to lysozymes. Nat Struct Mol Biol. 2005;12:270–3.

350. Keep NH, Ward JM, Cohen-Gonsaud M, Henderson B. Wake up! Peptidoglycan lysis and bacterial non-growth states. Trends Microbiol. 2006;14:271–6.

351. Turner D, Ackermann H, Kropinski AM, Lavigne R, Sutton JM, Reynolds DM. Comparative analysis of 37 *Acinetobacter* bacteriophages. Viruses. 2017;10:1–25.

352. Sazinas P, Redgwell T, Rihtman B, Grigonyte A, Michniewski S, Scanlan DJ, et al. Comparative Genomics of Bacteriophage of the Genus *Seuratvirus*. Genome Biol Evol [Internet]. 2018;10:72–6. Available from: http://academic.oup.com/gbe/article/10/1/72/4767719

353. Labrie SJ, Frois-Moniz K, Osburne MS, Kelly L, Roggensack SE, Sullivan MB, et al. Genomes of marine cyanopodoviruses reveal multiple origins of diversity. Environ Microbiol. 2013;15:1356–76.

354. Breitwieser FP, Lu J, Salzberg SL. A review of methods and databases for metagenomic classification and assembly. Brief Bioinform [Internet]. 2017;1–15. Available from: http://academic.oup.com/bib/article/doi/10.1093/bib/bbx120/4210288/A-review-of-methods-and-databases-for-metagenomic

355. Adams MJ, Antoniw JF, Fauquet CM. Molecular criteria for genus and species discrimination within the family Potyviridae. Arch Virol. 2005;150:459–79.

356. Casjens SR. Diversity among the tailed-bacteriophages that infect the *Enterobacteriaceae*. Res Microbiol. 2008;159:340–8.

357. Klyczek KK, Bonilla JA, Jacobs-Sera D, Adair TL, Afram P, Allen KG, et al. Tales of diversity: Genomic and morphological characteristics of forty-six *Arthrobacter* phages. PLoS One. 2017;12:1–30.

358. Grose JH, Jensen GL, Burnett SH, Breakwell DP. Correction to: Genomic comparison of 93 *Bacillus* phages reveals 12 clusters, 14 singletons and remarkable diversity. BMC Genomics [Internet]. 2014;15:855. Available from: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-855

359. 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.

360. O'Flaherty S, Coffey A, Edwards R, Meaney W, Fitzgerald GF, Ross RP. Genome of staphylococcal phage K: a new lineage of *Myoviridae* infecting gram-positive bacteria with a low G+C content. J Bacteriol. 2004;186:2862–71.

361. Kashani HH, Moniri R. Expression of recombinant pET22b-LysK-Cysteine/Histidine-Dependent Amidohydrolase/Peptidase bacteriophage therapeutic protein in *Escherichia coli* BL21 (DE3). Osong Public Heal Res Perspect. 2015;6:256–60.

362. Jarábková V, Tišáková L, Godány A. Phage endolysin: A way to understand a binding function of c-terminal domains a mini review. Nov Biotechnol Chim. 2015;14.

363. Vanhommerig E, Moons P, Pirici D, Lammens C, Hernalsteens JP, De Greve H, et al. Comparison of biofilm formation between major clonal lineages of methicillin resistant *Staphylococcus aureus*. PLoS One. 2014;9:e1044561.

364. Bhattacharya M, Wozniak DJ, Stoodley P, Hall-Stoodley L. Prevention and treatment of *Staphylococcus aureus* biofilms. Expert Rev Anti Infect Ther. 2015;13:1499–516.

365. Kaur S, Harjai K, Chhibber S. Methicillin-resistant *Staphylococcus aureus* phage plaque size enhancement using sublethal concentrations of antibiotics. Appl Environ Microbiol. 2012;78:8227–33.

366. Chibani-chennou S, Sidoti J, Bruttin A, Kutter E, Bru H. In vitro and in vivo bacteriolytic activities of *Escherichia coli* phages: implications for phage therapy. Society. 2004;48:2258–569.

367. Hyman P. Phages for phage therapy: Isolation, characterization, and host range breadth. Pharmaceuticals. 2019;12:35.

368. Ross A, Ward S, Hyman P. More is better: Selecting for broad host range bacteriophages. Front Microbiol. 2016;7:1–6.

369. Jensen KC, Hair BB, Wienclaw TM, Murdock MH, Hatch JB, Trent AT, et al. Isolation and host range of bacteriophage with lytic activity against methicillin-resistant

Staphylococcus aureus and potential use as a fomite decontaminant. PLoS One. 2015;10:e0131714.

370. Deurenberg RH, Rijnders MIA, Sebastian S, Welling MA, Beisser PS, Stobbering EE. The *Staphylococcus aureus*lineage-specific markers collagen adhesin and toxic shock syndrome toxin 1 distinguish multilocus sequence typing clonal complexes within spa clonal complexes. Diagn Microbiol Infect Dis. 2009;65:116–22.

371. Atshan SS, Shamsudin MN, Thian Lung LT, Sekawi Z, Ghaznavi-Rad E, Pei Pei C. Comparative characterisation of genotypically different clones of MRSA in the production of biofilms. J Biomed Biotechnol. 2012;417247.

372. Naicker PR, Karayem K, Hoek KGP, Harvey J, Wasserman E. Biofilm formation in invasive *Staphylococcus aureus* isolates is associated with the clonal lineage. Microb Pathog. 2016;90:41–9.

373. Parker RE, Laut C, Gaddy JA, Zadoks RN, Davies HD, Manning SD. Association between genotypic diversity and biofilm production in group B *Streptococcus*. BMC Microbiol. 2016;6.

374. Grohs P, Kitzis MD, Gutmann L. *in vitro* bactericidal activities of linezolid in combination with vancomycin, gentamicin, ciprofloxacin, fusidic acid, and rifampin against *Staphylococcus aureus*. Antimicrob Agents Chemother. 2003;47:418–20.

375. Scanlan PD, Hall AR, Burlinson P, Preston G, Buckling A. No effect of host-parasite co-evolution on host range expansion. J Evol Biol. 2013;27:51–63.

376. Torres-Barceló C, Hochberg ME. Evolutionary rationale for phages as complements of antibiotics. Trends Microbiol. 2016;24:249–56.

377. Pires DP, Melo LDR, Vilas Boas D, Sillankorva S, Azeredo J. Phage therapy as an alternative or complementary strategy to prevent and control biofilm-related infections. Curr Opin Microbiol [Internet]. Elsevier Ltd; 2017;39:48–56. Available from: http://dx.doi.org/10.1016/j.mib.2017.09.004

378. Benešík M, Nováček J, Janda L, Dopitová R, Pernisová M, Melková K, et al. Role of SH3b binding domain in a natural deletion mutant of *Kayvirus* endolysin LysF1 with a broad range of lytic activity. Virus Genes. 2018;54:130–9.

379. Burrowes BH, Molineux IJ, Fralick JA. Directed in vitro evolution of therapeutic bacteriophages: The Appelmans protocol. Viruses. 2019;11:241.

380. Kropinski AM. Phage therapy - Everything old is new again. Can J Infect Dis Med Microbiol. 2006;17.

381. Henry M, Lavigne R, Debarbieux L. Predicting in vivo efficacy of therapeutic bacteriophages used to treat pulmonary infections. Antimicrob Agents Chemother. 2013;57:5961–8.

382. Payne RJH, Jansen VAA. Understanding bacteriophage therapy as a density-dependent kinetic process. J Theor Biol. 2001;208:37–48.

383. Redfern J, Wallace J, Belkum A Van, Jaillard M, Whittard E, Ragupathy R, et al. Biofilm associated genotypes of multidrug-resistant *Pseudomonas aeruginosa*. bioRxiv. 2019;1–25.

384. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. Rev Antimicrob Resist [Internet]. 2016;84. Available from: https://amr-review.org

385. Schlievert PM, Strandberg KL, Lin YC, Peterson ML, Leung DYM. Secreted virulence factor comparison between methicillin-resistant and methicillin-sensitive *Staphylococcus aureus*, and its relevance to atopic dermatitis. J Allergy Clin Immunol. 2010;125:39–49.

386. Harrison JJ, Turner RJ, Marques LLR, Ceri H. A new understanding of these microbial communities is driving a revolution that may transform the science of microbiology. Am Sci. 2005;93:508–15.

387. Rohwer F, Maughan H, Youle M, Hisakawa N. Life in Our Phage World: A centennial field guide to the earth's most diverse inhabitants. illustrate. Wholon; 2014.

Appendix A

Table 6.1. *S. aureus* isolates list with their sequence type (ST) used throughout this study

Strain	ST	Strain	ST	Strain	ST
NOT 290	-	403.02	ST22	H118	ST28
15981	-	1091	ST22	CUBA 4030	ST30
BTN 2164	ST1	434.07	ST22	C390	ST31
HT 2001 - 254	ST1	370.07	ST22	H399	ST33
NL0118512	ST1	729192 April	ST22	C160	ST34
H462	ST1	W449 36	ST22	MRSA 252 ΦK MUTANT	ST36
CDC USA 800	ST5	HO 5322 0548 09	ST22	MRSA 252	ST36
BTN 2289	ST5	F869 56	ST22	NottnA2	ST36
FIN 61974	ST5	H431 62	ST22	H325	ST36
D10	ST5	WW1678/96	ST22	CDC 960758 USA 200	ST36
BK519	ST5	EMRSA15-90	ST22	BTN 766	ST36
NJ992 (CDC4)	ST5	ARI 26	ST22	EMRSA 16	ST36
AR110735	ST5	ARI 15	ST22	07.3841.N	ST36
A93 - 0066	ST5	ARI 7	ST22	07.2589.M	ST36
H157	ST5	ARI 31	ST22	06.9570.L	ST36
BTN 2242	ST5	SwedenAO9973	ST22	98.5806.F	ST36
C56	ST6	ARI 4	ST22	03.1791.F	ST36
EMRSA7	ST8	ARI 12	ST22	07.7206.Y	ST36
CDC USA 300	ST8	ARI 11	ST22	07.1227.Z	ST36
D137	ST8	ARI 10	ST22	07.1696.F	ST36
EMRSA 2	ST8	ARI 5	ST22	07.6659.K	ST36
E228	ST8	H65	ST22	07.2449.K	ST36
NL010548 - 1	ST8	C720	ST22	07.2496.L	ST36
99ST 2211	ST8	98.4823.X	ST22	97.2483.Hb	ST36
NOT 110	ST8	99ST18131	ST22	07.6636.Y	ST36
EMRSA 6	ST8	98/10618	ST22	07.2880.V	ST36
EMRSA 13	ST8	AR 0650 784	ST22	UK96/32010	ST36
D295	ST9	BTN 1626	ST22	NottmA	ST36
H169	ST9	NL011399-5	ST22	BTN 2172	ST36
D316	ST11	H182MRSA	ST22	FIN75916	ST36
D329	ST12	RH 0600 0061/09	ST22	BTN 1429	ST36
H117	ST12	HO 7230 0407/05	ST22	BTN 2292	ST36
H402	ST13	HO 5096 0412	ST22	H119 MRSA	ST36
C154	ST14	H914 91	ST22	H137	ST38
C357	ST15	C101	ST22	C253	ST40
H291	ST18	T277 06	ST22	C427	ST42
D17	ST20	M810 08	ST22	BTN 2306	ST45
723.07	ST22	HO 7374 0468/05	ST22	FIN 76167	ST45
1018.07	ST22	T505 30	ST22	BTN 2299	ST45
921.07	ST22	C49	ST23	C316	ST49
930.02	ST22	D279	ST25	H417	ST50

Table 6.1. Continued.

Strain	ST
C3	ST51
D49	ST53
D98	ST54
D97	ST55
D318	ST57
D508	ST58
D535	ST59
D551	ST59
D473	ST69
CDC USA 700	ST72
HT 2004 0991	ST80
HT 2002 - 0664	ST80
SWEDEN 8890/99	ST80
BK 1563	ST88
HT 2002 - 0635	ST93
HT 2001 - 634	ST93
H560	ST121
D139	ST145
FIN 62305	ST156
D22	ST182
CAN 6428 - 011	ST188
D470	ST207
NOT 116	ST227
NOT 161	ST227
GERMANY 131/98	ST228
WW 2594/97 - 2	ST229
CDC 16	ST231
99.3759.V	ST235
EMRSA 11	ST239
FFP200	ST239
EMRSA 4	ST239
91 - 4990	ST239
EMRSA 9	ST240
SWEDEN 408/99	ST246
FRA 97393	ST247
EMRSA 5	ST247
82 MRSA	ST247
KD 12168	ST250
EMRSA 8	ST250
12.2539.L	ST398
C7-t011	ST398

Strain	ST
h-RVC57276	ST398
11.1299.J	ST398
11.4910.K	ST398
11.5252.H	ST398
GKP136-53	ST398
C7 (P11)	ST398
42-57	ST398
11.5654.T	ST398
BVCA92	ST398
m-38-53	ST398
RV2007-13689-13	ST398
RV2007-06745-3 'A'	ST398
27969	ST398
m-mecA-17-57	ST398
12.2732.H	ST398
09.4620.V	ST398
12.2167.C	ST398
09.6440.M	ST398
C7 (P4)	ST398
11.3281.H	ST398
11.2530.K	ST398

Table 6.2. Top hits of genes sharing high structural homology to lytic enzyme proteins using the structure prediction tool HHpred.

Gene	Annotation	HHpred Top Hit	Protein ID	Similarity	E-value
group_413	Tail lysin	Peptidoglycan hydrolase	4Q4G_X	99.59	5.90E+17
isaA	ORF054	Membrane-bound lytic murein transglycosylase	4P0G_A	98.3	1.20E-08
group_11	ORF042	LysK / Hydrolase	4CSH_A	99.8	6.40E-24
group_10	Hypothetical protein	LysK / Hydrolase	4CSH_A	99.4	5.80E-17
group_217	Amidase	LysK	4CSH_A	98.9	6.40E-13
group_742	N-acetylmuramoyl-L-alanine amidase	Bifunctional autolysin	4KNK_A	98.5	1.00E-10
group_1366	Transglycosylase	Resuscitation-promoting factor / Hydrolase	2N5Z_A	98.66	3.20E-10
group_1636	Transglycosylase	Resuscitation-promoting factor / Hydrolase	4CGE_A	97.99	6.00E-08
group_329	Transglycosylase	Resuscitation-promoting factor / Hydrolase	4EMN_B	96.7	4.00E+10
group_23	Transglycosylase	Resuscitation-promoting factor / Hydrolase	4EMN_B	96.7	2.20E-05
Iss	N-acetylmuramoyl-L-alanine amidase	Endolysin / Hydrolase	2MK5_A	99.6	2.20E-20
group_298	ORF082	LysK / Hydrolase	4CSH_A	97	4.10E-06
group_1	ORF001	Morphogenesis protein / Hydrolase	3CT5_A	96.77	8.80E-05
group_4	Tape meaure protein / Tail lysin	Morphogenesis protein / Hydrolase	3CT5_A	98.8	5.00E-12
group_1374	Lysostaphin	Peptidoglycan protease	4CT3_D	99.2	3.70E-10
group_1238	N-acetylmuramoyl-L-alanine amidase domain-containing protein	CHAP domain / Amidase	2K3A_A	98	8.80E-09
group_414	Tail lysin	CHAP domain / Amidase	2K3A_A	98.8	5.00E-12
group_18	CHAP domain protein	CHAP domain / Amidase	2K3A_A	98.8	1.50E-11
group_19	Tail Lysin	CHAP domain / Amidase	2K3A_A	98.7	6.90E-12
group_20	Tail lysin	CHAP domain / Amidase	2K3A_A	99	2.90E-13
group_237	Tail lysin	N-acetylglucosaminidase	5AZM_A	99.3	3.00E-16

Table 6.2. continued.

Hhpred Second Hit	Protein ID	No. isolates	No. sequences
Resuscitation-promoting factor / Hydrolase	3NE0_A	11	11
Membrane-bound lytic murein transglycosylase	4CFP_B	22	22
Peptidoglycan protease	4CT3_D	32	32
Peptidoglycan protease	4CT3_D	1	1
Endolysin / Hydrolase	2MK5_A	11	11
Endolysin / Amidase	4OLS_D	11	11
Lysozyme-like sturcture	1XSF_A	1	1
Membrane-bound lytic murein transglycosylase	4P0G_A	1	1
Lysozyme	2VB1_A	11	11
Lysozyme	2VB1_A	15	15
Endolysin / Amidase	4OLS_D	39	39
Peptidoglycan protease	4CT3_D	19	19
Peptidoglycan endopeptidase	4Q4G_X	37	37
Peptidoglycan endopeptidase	4Q4G_X	12	12
LysK	4CSH_A	1	1
LysK	4CSH_A	1	1
LysK	4CSH_A	11	11
LysK	4CSH_A	7	7
LysK	4CSH_A	42	42
LysK	4CSH_A	1	1
Autolysin	4PIA_A	11	11