Identification and Characterization of Six Bacteriophages Capable of Infecting Extremely Antibiotic Resistant Strains of *Stenotrophomonas maltophilia*

by

Danielle Lynne Peters

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in
Microbiology and Biotechnology

Department of Biological Sciences

University of Alberta

Abstract

Infections caused by the ubiquitous Gram-negative bacillus *Stenotrophomonas maltophilia* are on the rise worldwide. Due to the impressive array of innate antibiotic resistance mechanisms employed by this bacterium, treatment of infections is exceedingly difficult; therefore, alternative treatment options are required. To combat these deadly infections, researchers around the world have shown a renewed interest in the use of phage therapy, or the clinical application of viruses parasitic to bacteria. To this end, six phages were isolated against three clinical isolates of *S. maltophilia* for characterization.

The suitability of each phage was assessed for therapeutic use by analyzing the complete genome sequences. The genome sizes range from 42 to 168 kilobase pairs in length, and many of the phages show a surprising amount of diversity from any phage characterized to date. Of the six phages isolated, three have a lytic lifestyle and do not encode virulence factors. Two of the lytic phages, DLP1 and DLP2, have extended host ranges beyond *S. maltophilia* as they are capable of infecting strains of *Pseudomonas aeruginosa*. The third lytic phage, DLP6, is a T4-like *Myoviridae* phage with a moderate host range. These phages would be suitable for inclusion in a phage cocktail due to their lytic lifestyles and host ranges.

In contrast to the lytic phages identified, the three temperate phages characterized would not be suitable for inclusion in a phage cocktail due to their ability to cause lysogenic conversion of their hosts. Phages DLP3 and DLP5 are identified as members of a novel genus, *Delepquintavirus*, due to their limited identity to other known phages. DLP3 lysogenization leads to an increase in erythromycin resistance and growth rate of its host strain. DLP5 is capable of lysogenizing its host as a phagemid and causes an increased growth rate during the lag and early exponential phase of growth. Phage DLP4 is capable of stable lysogeny even though no genes

encoding proteins involved in a temperate lifestyle were identified. Lysogenization by DLP4 leads to increased trimethoprim resistance and the expression of a virulence factor known to increase swarming in *Escherichia coli*. Although only three lytic phages were identified as good candidates for phage therapy, characterization of the temperate phages offered insight into the important role phages play in the evolution of *Stenotrophomonas maltophilia* with respect to antibiotic resistance and putative virulence factors.

Acknowledgements

The entirety of the work presented here would not have been possible without the love and support of an amazing group of people.

I would first like to thank Dr. Jonathan Dennis for his guidance, support, and faith in me. Throughout the many life-changing events I endured over the course of this work, you believed in my capabilities as a scientist and it encouraged me to push on. I cannot express how much that means to me. Thank you.

Thank you to all of the past and present Dennis and Raivio lab members for all of the stimulating discussions and fun group outings. To Dr. Fatima Kamal, Dr. Randi Guest, and Jaclyn McCutcheon - you all had a great influence on me as a scientist and helped me enjoy life in Edmonton more than I thought possible. A special thank you to Catherine Tays for being an amazing friend and brilliant colleague. You supported me and made me laugh when I didn't know it was possible.

Thank you to my committee members Dr. Stephan Pukatzki and Dr. Paul Stothard for your valuable insights and support throughout my graduate degree. Additionally, thank you Dr. Pukatzki, Dr. Lisa Stein, and Dr. Tracy Raivio for your engaging discussions and for kindly writing me reference letters.

Thank you to Alberta Innovates – Technology Futures, NSERC, and the Queen Elizabeth II awards which allowed me to follow my dream while paying off my entire student loan.

Thank you to my family spread across British Columbia for their unwavering belief in me and providing perspective when things went wrong: Deane, Meagan, Barb, Jim, Dad, Donell, Mom, and Terry.

Finally, thank you to Deane. The adventures we have shared together helped pull me though my graduate degree. You reminded me to go outside, have fun, and smell the roses. I could not have done this without your encouragement to always keep going. Thank you for teaching me how to eat an elephant.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA	amino acid
AB26	*Pseudomonas* phage vB_PaeS_SCH_Ab26
bp or kbp	base pair or kilobase pair
BSI	Blood stream infection
CANWARD	Canadian Ward Surveillance
CBCCRRR	Canadian *Burkholderia cepacia* complex Research and Referral Repository
CD-Search	conserved domain search
CDS	coding domain sequence
CF	cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CFU	colony forming units
Cm	chloramphenicol
˚C	Degrees Celsius
DNA	deoxyribonucleic acid
DUF	domain of unknown function
EPEC	Enteropathogenic *Escherichia coli*
ETEC	Enterotoxigenic *Escherichia coli*
erm	erythromycin ribosome methylase
g or mg or ng	grams or milligrams or nanograms
gDNA	genomic DNA
gp	gene product
h	hour
ICU	intensive care unit
ID	identity
IPATH	Innovative Phage Applications and Therapeutics
IS	insertion sequence

kV	kilovolt
l or ml or µl	liter or millilitre or microliter
LB	Lauria Bertani
min	minute
mL	milliliter
mM or M	millimolar or molar
MOI	multiplicity of infection
MUSCLE	MUltiple Sequence Comparison by Log Expectation
N/A	not applicable
NCBI	National Center for Biotechnology Information
NCF	non-cystic fibrosis
nm	nanometer
NSERC	Natural Sciences and Engineering Research Council
OD	optical density
ORF	open reading frame
PA73	*Pseudomonas aeruginosa* phage73
PCR	polymerase chain reaction
PFU	plaque forming units
PK25	*Pseudomonas* phage vB_Pae-Kakheti25
RBS	ribosome binding site
RFLP	restriction fragment length polymorphism
RNAP	RNA polymerase
RT	reverse transcriptase
s	seconds
Sec	secretion route
SM	suspension medium
Smc	*Stenotrophomonas maltophilia* complex
SNP	single nucleotide polymorphism

T2S	type II secretion
Tat	twin-arginine
T4P	type IV pilus
TMP/SMX	trimethoprim/sulfamethoxazole
Tp	trimethoprim
tRNA	transfer ribonucleic acid
Vsr	very short patch repair

Chapter 1

Introduction

Stenotrophomonas maltophilia complex (Smc)

Isolation and taxonomy

In 1943, J. L. Edwards isolated an unknown aerobic Gram-negative bacillus from a pleural fluid specimen which he subsequently named *'Bacterium bookeri'*[1]. Further study into this bacterium led to its re-classification as *Pseudomonas maltophilia* due to the growth and morphological similarities the strain shared with members of the *Pseudomonas* genus[1]. Additional strains of *P. maltophilia* were later identified which had been incorrectly classified as *Pseudomonas melanogena*[2], *Pseudomonas alcaligenes*[3], and *Alcaligenes faecalis*[4]. As the field of molecular biology progressed, further investigation into *P. maltophilia* strains with DNA and r-RNA hybridization techniques was possible.

Researchers found the rRNA cistrons of *P. maltophilia* were most similar to three strains of *Xanthomonas*[5]. This discovery led to the creation of the genus *Stenotrophomonas*, with *Stenotrophomonas maltophilia* the type species of the genus[6]. As of September 1st, 2019, there are 19 species of *Stenotrophomonas* in the NCBI taxonomy browser, though more species are likely to be identified due to the diversity of this genus. The species identified to date include: *S. maltophilia, Stenotrophomonas rhizophila*[7], *Stenotrophomonas nitritireducens*[8], *Stenotrophomonas acidaminiphila*[9], *Stenotrophomonas terrae*[10], *Stenotrophomonas humi*[10], *Stenotrophomonas koreensis*[11], *Stenotrophomonas chelatiphaga*[12], *Stenotrophomonas bentonitica*[13], *Stenotrophomonas daejeonensis*[14], *Stenotrophomonas detusculanense, Stenotrophomonas ginsengisoli*[15], *Stenotrophomonas indicatrix*[16], *Stenotrophomonas lactitubi*[16], *Stenotrophomonas indologenes, Stenotrophomonas panacihumi*[17], *Stenotrophomonas pavanii*[18], *Stenotrophomonas pictorum*[19], and *Stenotrophomonas tumulicola*[20]. Further phenotypic and genotypic studies of these species, as well as analysis of their ecological and metabolic diversity, has enabled some species-level differentiation, but researchers found these methods to be unreliable within the *Stenotrophomonas* genus. To remedy this, researchers compared the *gyrB* gene sequence, which encodes the beta subunit of DNA gyrase, for accurate interspecies differentiation and identification[21].

The *gyrB* results revealed large variability within *S. maltophilia* strains whereby strains classified as *S. maltophilia* were found to share > 99.0 % 16s rRNA sequence identity while

having less than 70 % pairwise identity to the genomic DNA of the CCUG 5866 type strain[21]. Additionally, several species closely related species to *S. maltophilia* such as *S. pavanii, S. africana, Pseudomonas geniculata, Pseudomonas hibisciola,* and *Pseudomonas beteli* were identified, and researchers suggested the group of bacteria should be referred to as the *S. maltophilia* complex (Smc) [21].

The following year, Rhee et al. (2013) published on the phylogenetic analysis of 118 Smc isolates from seven Korean hospitals. The data led to the identification of three groups, I to III[22]. Groups II and III cluster into a single clade with *S. pavanii* and *P. beteli,* while group I formed a clade with *P. geniculata* and *P. hibisciola.* The antimicrobial resistance rates for all Smc isolates were tested, and each group showed differing resistance levels, though resistance to the recommended *S. maltophilia* treatment of trimethoprim/sulfamethoxazole (TMP/SMX) was high (30.5 %) in all isolates. Resistance rates for isolates from groups II and III were significantly lower than isolates from group I; multi-drug resistance (MDR) was noted in 25 % of group II and III isolates, while 46 % of group I isolates were multidrug resistant[22].

A subsequent study using phylogenetic and population genetics on sequences from complete and draft genomes and 108 novel environmental isolates from Central Mexico were used to further investigate the Smc phylogeny[23]. At least five genospecies with significant differences were identified: *S. maltophilia* and Smc1 to Smc4. The only Smc lineage intrinsically MDR was identified as *S. maltophilia*, with all members expressing metallo-β-lactamases[23]. The genospecies identified also show habitat preferences, with *S. maltophilia* preferring contaminated sediments and the Smc1 and Smc2 genospecies preferring water columns of clean or moderately contaminated sites[23]. The authors mention further comparative and population genomic studies are required to resolve pending issues with the *S. maltophilia* sub-lineages.

Morphology and culture characteristics

Stenotrophomonas maltophilia is a non-sporulating Gram-negative bacillus which measures 0.5 to 1.5 μm long[24]. *S. maltophilia* is an obligate aerobe which cannot grow at temperatures below 5 ˚C or higher than 42 ˚C. The optimal growth temperature for clinical strains is around 35 ˚C[25]. In rich media, the doubling time for most *S. maltophilia* strains is between 20 to 40 minutes during log phase[26]. Overnight cultures can reach densities of > 10^9 cell forming units (CFU)/ml; though, incubation past 20-24 h can result in significant death phase[26].

This may be due to induction of the lytic cycle in lysogenic strains. Most strains can grow on a wide variety of media, though due to the salt sensitivity of *S. maltophilia,* researchers are cautioned when using high-salt media formulations[26]. Personal observations have shown high salt concentrations can lead to the induction of prophages in lysogenic strains. Incubation of clinical isolates at 30 ˚C for 16 to 36 h is required for the formation of single colonies (personal observations). Single colonies of *S. maltophilia* appear smooth and white to pale yellow in color, and some strains glisten[24]. The bacterium is motile due to the presence of several polar flagella[27-29]. Strains of *S. maltophilia* have also been shown to be capable of flagella-independent translocation in the presence of extracellular fatty acids[28].

Biotechnology applications

The genus name is derived from the Greek ‘stenos,’ ‘trophus,’ and ‘monas,’ which combine to mean, ‘narrow one who feeds,’ and ‘unit,’ respectively. The genus name was chosen due to the perceived limited nutritional range of the bacterium[6], though several studies have since shown the impressive metabolic versatility and intraspecies heterogeneity of *Stenotrophomonas*[30-34].

The metabolic diversities of *S. maltophilia*, which are highlighted below, have inspired researchers to investigate their use in bioremediation. In 2017, Salah-Tazdait et al. published a paper on *S. maltophilia* isolates which were found to be capable of using the insecticides fenitrothion and malathion as their sole carbon source; and in the case of fenitrothion, sole nitrogen source[35]. These strains could be used in the bioremediation of soils contaminated with insecticides. Further, a biosurfactant-producing strain of *S. maltophilia* was identified and investigated for biodegradation of diesel oil and used engine oil. The strain was capable of growing on mineral salts medium supplemented with diesel oil and used engine oil as the sole carbon and energy source[36]. This strain could be used in the bioremediation of contaminated waterways or soils due to spilled diesel or engine oil.

Further highlighting the metabolic plasticity of *S. maltophilia,* researchers have isolated strains from sites contaminated with heavy metals. Two groups have each isolated a strain capable of reducing the known carcinogen and mutagen, hexavalent chromium Cr(VI)[37,38]. Hexavalent chromium is used in wood preservation, anti-corrosion products, textile dyes, and a variety of other niche uses. Contamination of soils and waterways is being identified as a major

health concern due to the contamination of municipal water sources. One strain showed resistance up to 400 mg/ml[38], well above the contamination levels identified in municipal water sources. The high Cr(VI) concentration resistance and high Cr(VI) reducing ability of the strains make them suitable candidates for bioremediation.

Further bioremediation and phytoremediation possibilities with *S. maltophilia* includes decontamination of copper from waterways and shorelines. Researchers in India have isolated multiple *Stenotrophomonas* sp., including one *S. maltophilia* strain, from a site heavily contaminated with brass industry effluent[39]. These strains were found in the rhizosphere of the phytoremediation plant *Cynodon dactylon* and showed high resistance to copper. The *S. maltophilia* strain was copper-tolerant and featured many plant growth-promoting mechanisms such as the siderophore secretion, phosphate solubilization, and the production of 1-Aminocyclopropane-1-carboxylate deaminase (ACC)[39]. ACC is reported to protect the plant from heavy metal toxicity, resulting in the formation of longer and denser roots in contaminated sites[40].

Members of the Smc also play important ecological roles in nitrogen and sulphur cycles[41-43], and unlike the closely related phytopathogenic genera *Xanthomonas* and *Xylella, Stenotrophomonas* species are not phytopathogenic[44]. Members of the complex are often found in close association with plant rhizospheres where they have been shown to actively promote plant growth though the secretion of growth promoting compounds[44]. Researchers in Korea isolated diazotrophic growth promoting bacteria from the rhizosphere of agricultural crops for their potential use as biofertilizers[45]. Two strains of *S. maltophilia* were identified, and the strains were shown to have high nitrogenase activity while also secreting high levels of the plant promoting hormone indole-3-acetic acid (IAA) [37], making them potential biofertilizer candidates[43,46].

Enzymes produced by *S. maltophilia* complex members are being researched for their biotechnology applications. For example, the environmental *S. maltophilia* Psi-1 strain was isolated from a Greek sludge sample and was found to encode a novel bacterial lipase (LipSm; lipase family XIX)[47]. Microbial lipases catalyze a broad array of reactions, making them enzymes of considerable biotechnological interest[48]. The LipSm was found to be alkaliphilic, thermostable, and 3D modelling revealed a lid-structure which enables LipSm to act without the need for interfacial activation with small substrates[47].

Collectively, these research findings are just a glimpse of the biotechnological applications possible with members of the Smc. The vast metabolic diversities observed in the Smc enables these bacteria to have highly beneficial uses in bioremediation, such as heavy metal detoxification of soils and waterways to the degradation of insecticides and volatile organic compounds such as benzene. Additionally, the plant growth promotion associated with Smc members in the rhizosphere is being exploited though their use as biofertilizers due to their ability to fix nitrogen, produce growth promoting plant hormones, and protect the plant roots from phytopathogens. Thus, the *Stenotrophomonas* complex members have great potential for use in the biotechnology industry, though their widespread use without proper precautions is troublesome due to their ability to cause disease in humans. Some researchers have advocated the use of *S. rhizophila* strains in biotechnological applications instead of *S. maltophilia* because *S. rhizophila* does not have the associated human health risks, mainly due to the inability of the bacterium to grow at 37 ˚C[45].

Clinical significance

The *Stenotrophomonas maltophlia* complex is a global emerging Gram-negative multi-drug resistant group of organisms most commonly associated with pneumonia and bacteremia in immunocompromised patients. *S. maltophilia* has also been identified as the cause of soft tissue infections, osteomyelitis, meningitis, endocarditis, otitis, and scleritis[24]. There are numerous risk factors associated with *S. maltophilia* infections in the general population which include[49]: malignancy, human immunodeficiency virus, cystic fibrosis, intravenous drug use, surgical and accidental trauma, prolonged hospitalization, admission to the intensive care unit, mechanical ventilation, indwelling vascular and urinary catheters, corticosteroids, immunosuppressive therapy, treatment with broad-spectrum antibiotics, travel to hospital by air, gastrointestinal tract colonization/ mucositis, and hematopoietic stem cell transplantation. The numerous risk factors associated with this bacterium and increased surveillance programs for infections caused by this organism has led to an increase in the infection prevalence. Infection prevalence for this bacterium in the general population has increased from the 1997-2003 values of 0.8 – 1.4 % prevalence to 1.3-1.68 % when measured between 2007-2012[50]. These findings are consistent with both Canadian Ward Surveillance Studies (CANWARD) completed between 2007-2012[51,52]. Additionally, Naidu et al. published a study in 2012 from the University of Alberta

(UofA) tertiary care teaching hospital on the rates of *S. maltophilia* blood stream infections (BSI) per 1000 admissions was tracked over an 11 year period[53]. The results show admission rates varied from 0.04 to 0.22, and more than half of the *S. maltophilia* infections were classified as nosocomial. The mortality rates attributed to *S. maltophilia* bacteremia at the UofA hospital were 16.7 %, and the crude mortality rate of patients who had been infected with *S. maltophilia* was 25 %.

Older isolation and characterization reports on *S. maltophilia* infections suggested this bacterium had limited pathogenicity when infecting immunocompromised individuals; though, this is proving to be false with increasing reports of significant case/fatality ratios in severely debilitated or immunosuppressed patients[54]. This is highlighted by a review published by Velazquez-Acosta et al. in 2018 on the clinical outcomes of cancer patients with BSIs and pneumonia caused by *S. maltophilia* infections over a 16-year period at a tertiary-care oncology hospital in Mexico City [55]. Within the first month, 31.6 % of the patients had died; 70 % of deaths were due to pneumonia, and 30 % were due to BSI. All of the pneumonia cases were nosocomial, with 60.6 % of the cases ventilator-associated. Most BSI cases were from ambulatory patients with indwelling catheters (67.4 %), followed by nosocomial catheters (21.0 %), and secondary infection site BSI (11.6 %). These research findings were supported by an additional study by Schwab et al. (2018) on BSI mortality rates associated with 937 German intensive care units (ICU) from 4,556,360 patients[56]. The researchers found *S. maltophilia* infections had the highest mortality rates (28.4 %), followed next by non-albicans-Candida *spp.* (27.1 %), and *Pseudomonas aeruginosa* (25.8 %). The latest reviews on *S. maltophilia* infections highlight the virulence and pathogenicity of clinical isolates.

Transmissibility

Stenotrophomonas is highly capable of withstanding extremely inhospitable environments, which makes this bacterium troublesome in hospital settings. For example, nosocomial isolation sources for this bacterium include ultra-pure water, chlorhexidine-cetrimide topical antiseptic solution, hemodialysis water, nebulizers, and hand-washing soap[24,57]. Further, *S. maltophilia* has been shown to be able to tolerate hypochlorite cleaners, triclosan, sodium dodecyl sulfate, hospital antiseptic solution, and antiseptics containing quaternary ammonium compounds[24]. The bacterium is also extremely well adapted to handle periods of prolonged nutritional stress by forming ultramicrocells (UMC), which are capable of passing through a 0.2

μm filter[58]. The UMC pose an increased risk to patients because they can pass though the filters in point of use (POU) filtration units found in hospitals, which could have an aerosolization effect on the UMC as a patient uses the shower[24]. A spike (54 % from 2.6 %) in *S. maltophilia* infections was noted over seven days at Chosun University Hospital from bronchoalveolar lavage (BAL) specimens. Investigators took environmental samples and identified a contaminated fiberoptic bronchoscope suction channel that was the source of the outbreak. It was determined that inadequate cleaning and disinfection led to the outbreak[59].

Research into cough-generated aerosols from cystic fibrosis patients showed the production of respiratory particles containing viable *S. maltophilia*[60]. In the study, a cough aerosol sampling system was used to obtain aerosolized droplets from CF patients though forced coughing. *S. maltophilia* was cultured from four patients, two of which had tested sputum culture negative. An additional study was published in 2014 on cough-generated aerosols from CF patients to analyze the distance and persistence in the air of aerosolized *P. aeruginosa* and other CF pathogens[61]. *S. maltophilia* was also isolated from two out of 19 CF patients. Although data for the persistence and spread of *S. maltophilia* in the air was not provided, the *P. aeruginosa* isolates were found to travel up to 4 m and persist in the air for 45 min. Together, the results show cough-generated aerosols are a potential source of transmission for *S. maltophilia*, though additional research is required to determine the concentration and size of respiratory particles required to cause infection in susceptible individuals.

While being frequently recovered in clinical settings, multi-drug-resistant (MDR) *S. maltophilia* strains are also ubiquitous in non-clinical environments. Environmental isolates have been recovered from a variety of sources such as soil[62], water (lakes, rivers, ultrapure and tap)[24,63,64], plants[65], and food sources [24,66,67]. These sources can contribute to the community-acquired infections, which appear to be on the rise[68,69].

Virulence and Pathogenicity

Although *S. maltophilia* is not considered a highly virulent pathogen in healthy individuals, the bacterium is an important opportunistic pathogen. Due to a wide array of pathogenicity factors, the bacterium can be difficult to clear once an infection has been established. An increased mutation rate has been associated with clinical isolates compared to environmental isolates, suggesting clinical isolates are able to adapt to their local environment within the human body[30]. Mutation frequency studies from 174 clinical and nonclinical *S.*

maltophilia isolates revealed hypomutators were common in environmental strains (58.3 %), but hypermutators were only found in clinical isolates[70]. The results suggest clinical environments may select bacterial populations with high mutation frequencies due to their ability to quickly adapt to changes in their environment. Mutations in the *mutS* gene of the mismatch repair system were commonly responsible for the hypermutator phenotypes observed[70].

S. maltophilia pathogenicity is closely regulated based on the iron status of the cells[71]. Lactoferrin in the human lung sequesters iron away from microbial pathogens[24], though these low iron conditions may exacerbate *S. maltophilia* infections. Low iron conditions have been shown to cause an increase in many pathogenic cellular activities of *S. maltophilia* such as biofilm formation, outer membrane protein expression, DSF production, oxidative stress response, and virulence[71]. Researchers have also identified a pleiotropic gene, *ax21,* which encodes an outer membrane protein involved in biofilm formation, motility, tolerance to tobramycin and virulence in *Galleria mellonella* [72]. The researchers also noted additional changes in the expression of other genes involved in virulence and antibiotic resistance following deletion of *ax21*. Although *S. maltophilia* is not considered a highly pathogenic bacteria, this opportunistic pathogen does encode many pathogenicity factors which result in high mortality rates in immunocompromised patients. *S. maltophilia* pathogenicity mechanisms include biofilm formation, lipopolysaccharide, hydrolytic enzymes, two-type II protein secretion systems, host cell invasion, and a diffusible signal factor system (DSF)[24].

Hydrolytic enzymes

Sequencing of the *S. maltophilia* K279a genome revealed many extracellular enzymes encoded such as fibrolysin, lipases, esterase, DNase, RNase, and proteases. [73]. A study by Figueirêdo et al. in 2006 using *S. maltophilia* clinical isolates revealed their cytotoxic activity[74]. Clinical *S. maltophilia* culture supernatants applied to Hep-2 (human larynx epidermoid carcinoma) cells resulted in intensive rounding, loss of intercellular junctions, and membrane blebbing followed by cell death within 24 h[74]. Additionally, when Vero (African green monkey) and HeLa (human cervix) cells were exposed to the culture supernatants, the cytotoxic effects observed included vigorous endocytosis and cell aggregation. The use of protease inhibitors did not prevent cell cytotoxicity. Cell-free hemolytic activities were also observed with the clinical isolate supernatants. Additional virulence factors documented in the study included protease,

lipase, and lecithinase[74]. The hydrolytic enzymes identified in the *S. maltophilia* clinical isolates highlight the need for further studies focusing the role of these virulence factors in the establishment and persistence of human infections.

Lipopolysaccharide

Stenotrophomonas maltophilia has lipopolysaccharide (LPS) which contains lipid A, core oligosaccharide, and O-antigen. Structural studies have been carried out on the O-specific polysaccharide, the core oligosaccharide, and the lipid A from the LPS. The core oligosaccharide of strain NCTC 10257 has been shown to contain residues of D-glucose, D-mannose, D-galactosamine, D-galacturonic acid, and a 3-deoxyoctulosonic acid[75]. Additionally, the lipid A of this strain is constructed from phosphorylated glucosamine residues featuring *O*- and *N*- fatty acyl substituents. Components of the O-antigen have been identified as fucose, xylose, rhamnose, and glucose[76,77]. Bacterial cell adhesion to surfaces by charged LPS has been reported in *S. maltophilia*[24]. Two operons, *xanAB* and *rmlBACD*, are required for O-antigen and core region biosynthesis[78]. Biofilm production experiments on polystyrene surfaces using *rmlA, rmlC,* and *xanB* mutants showed decreased biofilm production from these strains compared to wild type control[78]. Additionally, the *rlmA,* and *rlmC* mutants produced significantly more biofilm when grown on a glass surface compared to the wild-type and *xanB* mutant. These results show the involvement of *S. maltophilia* LPS in surface attachment.

Further investigation into the role of LPS in surface attachment led to the identification of the *spgM* gene of *S. maltophilia.* The *spgM* gene encodes a bifunctional enzyme with phosphoglucomutase and phosphomannomutase activities which is important for assembly of the O-polysaccharide chain[77,79]. The *spgM* mutant strains also have lower yields of high-molecular-weight O-antigen compared to their parental strains[77,79]. Biofilm experiments with a *spgM* mutant showed increased biofilm formation on polyvinyl chloride, polystyrene, and borosilicate glass when compared to the parental strain[79,80]. The authors of the *spgM* biofilm quantification studies suggested the incomplete LPS of the mutant enabled the cells to adhere to the plastic or glass surfaces with exposed cell surfaces which are typically concealed by LPS[79,80].

Although the *spgM* mutants were found to have increased biofilm formation on plastic and glass surfaces, researchers wanted to determine how readily these mutant strains can establish infections compared to the wild type strains. A study of *S. maltophilia* LPS influence

on colonization and virulence in a rat lung model using a *spgM* mutant was completed[77]. The results showed decreased colonization of the knockout *spgM* strain compared to the parental strain. The authors mentioned the *spgM* mutant was found to be completely avirulent in the animal model. The *spgM* mutant complemented with *spgM* in trans restored the mutants ability to colonize rat lungs which provides evidence that full-length LPS is required for colonization. Additionally, the authors found the *spgM* mutant was susceptible to complement-mediated cell killing which was not observed with the wild-type or the complemented mutant; thus, LPS is an important virulence factor involved in *S. maltophilia* infection.

Biofilm formation

The ability of a bacterium to form biofilms is recognized as an important virulence trait which contributes to disease progression by providing protection against antibiotics and the hosts' immune system. *S. maltophilia* can colonize and form biofilms on many surfaces such as glass, plastics, medical devices, implantable products, and lung epithelial cells[78,81-83]. A study focusing on the biofilm formation of cystic fibrosis (CF) -derived IB3-1 bronchial epithelial cell monolayers revealed all isolates used in the study were able to form biofilms of varying degrees on the bronchial epithelial cells[83]. Type-1 fimbriae genes have only been identified in clinical *S. maltophilia* isolates, which suggests they may play a role in the colonization of infected individuals [29]. *In vitro* tissue culture assays have indicated the *S. maltophilia* fimbriae 1 (SMF-1) protein is important for their adherence to eukaryotic cells and glass[84]. Using anti-SMF-1 antibodies, *S. maltophilia* adherence to eukaryotic cells and glass was inhibited if the antibodies were applied during the early stages of infection[84]. The ability of isolates to form biofilms, and invade host cells, can contribute to the prolonged infections noted in hematology and oncology patients, despite aggressive antibiotic treatments[85].

Invasion of host cells

Stenotrophomonas maltophilia isolates can invade human bronchial epithelial cells[83,86]. Transmission electron microscopy experiments found cystic fibrosis (CF) and non-cystic fibrosis (NCF) isolates show both types of isolates equally adhere to intercellular junctions and invade human bronchial epithelial cells[24]. Invasion of transformed human bronchial epithelial 16 HBE14o- cells by CF and NCF isolates was found to be 0.45% and 0.40%, respectively[86]. In

contrast, invasion rates of CF IB3-1 bronchial cells by CF isolates varied from 0.01 to 4.94%[83]. The high susceptibility of CF IB3-1 cells could be due to lack of acidification in intracellular organelles due to the mutated cystic fibrosis transmembrane conductance regulator (CFTR) protein, allowing intracellular *S. maltophilia* to thrive in these vacuoles[87]. This scenario has been confirmed though observations of CF and NCF isolates inside of host epithelial cells[83]. Results obtained from the observations suggest *S. maltophilia* may be able to divide inside of membrane-bound endocytic vacuoles. *Stenotrophomonas* invasion of host cells could provide protection from the immune system and antibiotics, which could lead to the establishment of chronic infections.

Type II protein secretion systems

The type II secretion (T2S) system is used by Gram-negative bacteria to secrete proteins into the extracellular milieu or host organisms[88]. T2S is a two-step process which transports specific proteins across the inner membrane into the periplasm using either the secretion route (Sec) [89] or twin-arginine translocation (Tat)[90] pathway to be transported across the outer membrane by a complex of proteins which make up the T2S system[91]. Two T2S systems, Xps and Gsp, have been studied from the clinical *S. maltophilia* isolate K279a [92,93]. The Gsp T2S was cryptic in the experiment, though the researchers speculated it might be functional under different growth conditions[92]. Effects of the Xps T2S on A549 human lung epithelial cell line indicate virulence though cell rounding, actin rearrangement, cell detachment (3 h), and cell death after 24 hours. Mutation studies of the Xps T2S revealed it mediates the secretion of at least seven proteins [92,93].

Further investigation into the proteins secreted by the Xps T2S led to the discovery of the serine proteases StmPr1 – StmPr3[93,94]. The StmP1 protease was found to have proteolytic activities against collagen, fibrinogen, and fibronectin. Due to the contribution of the T2S to the virulence of many human pathogens such as *Escherichia coli*, *P. aeruginosa*, and *Vibrio cholerae*, more research on both type II secretion systems of *S. maltophilia* is necessary. The expression of the T2S of *S. maltophilia* may be regulated though diffusible signal factors, as an absence of these factors was shown to result in a decrease of extracellular proteases[24].

Diffusible signal factors

The closely related plant pathogen, *Xanthomonas campestris*, was found to use cis-Δ2-11-methyl-dodecenoic acid as a diffusible signal factor (DSF) to control virulence factor synthesis and virulence in plants[95]. Due to the close phylogenetic relationship of *S. maltophilia* to *X. campestris*, researchers hypothesised *S. maltophilia* may also us a DSF-dependent signalling system. Comparison of a K279a *rpfF* (regulation of pathogenicity factors) mutant to wild type revealed the gene is essential for DSF synthesis. Deletion of *rpfF* revealed its' pleiotropic nature as changes in virulence, biofilm formation, and motility were observed. Adding exogenous DSF to the mutant reversed the phenotypic effects[96].

The secretion of DSF may not be solely for the benefit of *S. maltophilia*; it may also be used for cross-species cell signaling during co-infections[95]. Research has demonstrated DSF of *S. maltophilia* is recognized by *Pseudomonas aeruginosa* in co-cultures[95]. Researchers discovered the biofilm architecture of a *P. aeruginosa* monoculture, or a co-culture with an *S. maltophilia rsfF*- mutant, is flat in appearance. When *P. aeruginosa* is co-cultured with *S. maltophilia* producing DSF, or exogenous *S. maltophilia* DSF is added, the biofilm architecture becomes filamentous[95]. *P. aeruginosa* also showed decreased susceptibility to polymyxin B and E in the presence of *S. maltophilia* DSF. It is evident that DSF is very important in the regulation of virulence in *S. maltophilia* and has also been implicated in inducing virulence factors of other pathogenic bacterial genera such as *Xanthomonas* and *Pseudomonas*.

Antibiotic resistance

Innate antibiotic resistance in Smc infections is a major contributing factor to treatment failure. Patients who receive the wrong antibiotic at initial infection diagnosis have an increased risk of mortality compared to patients who received appropriate antibiotic treatment from the start[55]. Treatment of infections can be further complicated by the emergence of mutants with pleiotropic resistance[54]. Several mechanisms have been identified which contribute to the high levels of innate and acquired antibiotic resistance. Intrinsic resistance is attributed to mechanisms such as reduced membrane permeability[97], chromosomally encoded multidrug efflux pumps, and chromosomally encoded Qnr pentapeptide repeat proteins [98-100]. Additionally, antibiotic-inactivating enzymes such as L1/L2 β-lactamases[101-103] and aminoglycoside- inactivating enzymes[104,105] have been identified in many *S. maltophilia* isolates. The drug of choice to treat *S.*

maltophilia infections continues to be trimethoprim/sulfamethoxazole (TMP/SMX)[106], as resistance rates are generally less than 10 %[107]. However, higher rates of TMP/SMX resistance, ranging from 16-78.8 %, have been noted in patients with cancer, cystic fibrosis, and in several countries such as Korea, Japan, Spain, Mexico, Saudi Arabia, Turkey, and Canada[50,108]. Thus, research into alternative treatment options is vital due to the increase in resistance to TMP/SMX noted with specific comorbidities and geographic locations.

To shed light on the various molecular mechanisms the Smc employs against antibiotics, a short overview of these mechanisms is required. The infamous innate resistance *S. maltophilia* has against β-lactams stems from the inducible chromosomally encoded L1 and L2 β-lactamases, and TEM-2 from a Tn1-like transposon[109]. The L1 protein is a molecular class B Zn2+-dependent metallo-β-lactamase which hydrolyzes virtually all classes of β-lactams, including penicillins, cephalosporins, and carbapenems[110]. The L2 protein is a molecular class A clavulanic acid-sensitive cephalosporinase[50]. Both L1 and L2 are regulated by AmpR, a transcriptional regulator located upstream of L2[101].

Similar to β-lactam resistance, the majority of trimethoprim/sulfamethoxazole (TMP/SMX) resistance is due to the presence of a *sul* gene instead of overexpression of efflux pumps. The *sul* gene, encoding a sulfonamide resistance protein, has been identified on class 1 integrons (*sul1*)[111,112] and insertion sequence common region (ISCR) elements (*sul2*)[113]. In addition to the *sul* genes, a *dfrA* gene encoding for dihydrofolate reductase associated with class 1 integrons has also been shown to cause high levels of resistance against TMP/SMX[114]. The efflux pumps SmeDEF, SmeOP-TolCsm, and SmeYZ, which will be briefly discussed below, have also been implicated in resistance to TMP/SMX; though, not to the same degree as the *sul* genes[115-117].

Efflux pumps are used by bacteria to transport toxic substances, such as antibiotics, from the cell into the environment. There are five families of efflux pumps which includes the resistance- nodulation-cell-division (RND) family, the major facilitator superfamily (MFS), the ATP binding cassette (ABC) family, the small multidrug resistance (SMR) family, and the multidrug and toxic compound extrusion (MATE) family[118]. There have been at least ten multidrug efflux pumps identified in *S. maltophilia* strains, with the most abundant type of efflux pumps belonging to the RND family[50]. These efflux pumps have been identified as SmeABC, SmeDEF, SmeIJK, SmeOP-TolCsm, SmeYZ, and SmeVWX [116,119-124]. Together, the RND efflux

pumps have been shown to protect *S. maltophilia* against quinolones, aminoglycosides, tetracycline, macrolides, chloramphenicol, TMP/SMX, levofloxacin, minocycline, nalidixic acid, doxycycline, and novobiocin[50].

The ABC family of efflux pumps encoded by *S. maltophilia* members includes SmrA and MacABCsm which provide resistance against fluoroquinolones and tetracycline (SmrA), and aminoglycosides, macrolides and polymixins (MacABCsm)[50,125]. MacABCsm efflux pump also plays an important role in regulating envelope and oxidative stress, as well as biofilm formation[125]. Two remaining efflux pumps identified in *S. maltophilia* are ErmCABsm and FuaABC. The ErmCABsm efflux pump belongs to the MFS family, providing resistance to hydrophobic compounds such as nalidixic acid and erythromycin[126], while FuaABC is a novel fusaric acid tripartite efflux pump which is induced by fusaric acid and may establish a new subfamily of tripartite efflux pumps[127].

Unlike most bacteria, *S. maltophilia* resistance to quinolones in clinical isolates is not often due to mutations in topoisomerases and gyrases; instead, resistance is commonly attributed to efflux pumps and a chromosomally encoded *qnr* gene[128]. The chromosomally encoded Qnr pentapeptide repeat protein provides low level resistance to quinolones[98-100], while overexpression of the multidrug efflux pumps SmeDEF and SmeVWX was found to be the most prevalent cause of quinolone resistance in *S. maltophilia* isolates[121,122].

Resistance to aminoglycosides, in contrast to quinolones, is primarily due to aminoglycoside-modifying enzymes. Three transferase enzymes responsible for resistance identified to date include a novel AAC(6')-lak (aminoglycoside acetyltransferase) from a Nepal MDR strain[129], AAC(6')-Iz (aminoglycoside acetyltransferase)[130], and APH(3')-IIc (aminoglycoside phosphotransferase)[104]. Additionally, efflux pumps have also been implicated in *S. maltophilia* aminoglycoside resistance. Overexpression mutants in the SmeABC, SmeOP-TolCsm, SmeYZ, or MacABCsm efflux pumps were shown to be responsible for high levels of aminoglycoside resistance[50], though the most common cause of resistance to aminoglycosides still remains the modifying enzymes discussed above.

Phage Therapy

Principles

Due to the high levels of innate antibiotic resistance found within *S. maltophilia* isolates, alternative treatment options are required. The use of bacteriophages, or viruses lytic towards bacteria, are being investigated as a viable option. This type of therapy would require the isolation and characterization of phages for their use to as a prophylactic or to treat infections. The procedure would involve the production of a high-titer, endotoxin-free phage stock. Ideally a cocktail of multiple phages targeting numerous bacterial strains with different phage-receptors would be developed in order to expand the host range of the cocktail and protect against receptor-mutation mediated resistance[131,132]. Once administered, the phages target their specific host, infecting and lysing the cell to release phage progeny into the surrounding environment. Depending on the phage, a single infection cycle can generate over 100 new virions per cell. These newly produced phages would then go on to infect other infecting bacterial cells until the pathogen is eventually cleared from the patient. Once there is no host remaining for the phage, the patients' immune system would eventually clear the phages from patients body[133]. Outlined below are several advantages to using phage therapy over traditional antibiotics[131,132,134]:

1. Bacteriophages have a narrow host range, often only infecting a few strains within a species or a few species within a genus. This enables targeted treatment for the affected individual. This spares the patients microbiome and decreases the risks associated with antibiotic consumption, such as the establishment of secondary infections like *Clostridium difficile*. It should be mentioned; the narrow host range of bacteriophages is also seen as a con to some researchers[132].
2. Bacteriophages are self-replicating; thus, *in situ* replication can significantly increase phage abundance at the site of infection exactly where it is needed. This is different from antibiotics which travel throughout the body and do not increase at the site of infection.
3. The high abundance of phage in the environment makes them easy to isolate, unlike antibiotics which requires rigorous testing to ensure the compound is not toxic and has acceptable side effects.
4. Phages are easy to manipulate though mutagenesis, evolution, and other molecular approaches to create therapeutically enhanced phages which perform better *in vivo*.

5. Phages are virulent against antibiotic resistant bacteria. Some phages even use virulence factors as receptors and mutations in the receptors leads to less virulent strains. These are referred to as anti-virulence phages.
6. No side effects have been reported during or after phage application, unlike antibiotics which can cause a variety of symptoms in patients ranging from minor complications to death if the patient is allergic.
7. Bacteriophage biology is well understood; therefore, the development of new therapeutic phages will require far less time and money. This is in stark contrast to novel antibiotics which requires millions of dollars and several years of study to ensure safety and efficacy.

Research addressing phage limitations

Increased interest in phage therapy has encouraged researchers to tackle some fundamental questions about the use of phages in a therapeutic sense – do phage pharmacokinetic properties change in the presence of human cells? How can phage cocktails be optimized to better protect against receptor-mutants? Can phage cocktails be enhanced for their use against intracellular pathogens? How can phages be protected against phage neutralizing antibodies during the course of treatment? Recently, researchers have begun to address these important questions though innovative approaches.

Investigation into the interactions of phage and bacteria in the presence of a human cell-line was conducted using the human pathogen *Clostridium difficile* and the phiCDH51 phage[135]. The *in vitro* experiments used the human colon tumorigenic cell line HT-29 to simulate the colon environment where *C. difficile* infection occurs. The results showed phages can reduce bacterial cell numbers more effectively in the presence of the HT-29 cell line than without. This observation was attributed to phage adsorbing to the colon cells, which may promote interactions between phage and bacteria. No toxic effects were observed with phage treatment alone, and bacterial lysis did not result in cytotoxic effects or toxin release. These results provide greater insight into phage-bacteria interactions in the presence of human cells, and further experiments by another group using a colon model have supported these findings. The researchers found *C. difficile* CFU rates dropped below the detectable limit with phage treatment while preventing toxin production which is often observed with antibiotic treatments[136]. These results help support

the use of phages in therapy as they show increased phage activity in the presence of mammalian cells and decreased *C. difficile* CFU rates without the production of toxins. These studies both used single phages, which is ill-advised in a therapeutic setting due to the possibility of receptor-mutation mediated resistance; thus, additional experiments using phage cocktails should be completed using the same methodologies.

To prevent against the development of phage-resistant receptor mutants, researchers have developed a protocol to isolate 'guard' phages for their inclusion in a phage cocktail[137]. To develop the phage 'guards', a susceptible *Escherichia coli* strain was exposed to the phage strain JDP1. Resistant colonies with large and small colony sizes developed, so both colony morphologies were picked and named Rb and Rs respectively. These isolates were then used to isolate two additional phages, RSP and RBP. The RSP and RBP phages were mixed with JDP1 to form a cocktail, with the two novel phages referred to as 'guard' phages because they kill JDP1 resistant mutants as they arise during the course of treatment. The researchers found the guard-killer phage cocktail increased bactericidal activity *in vivo* (mouse model) and inhibited the formation of phage resistant-strains. This approach to developing phage cocktails is handy as it is a generalizable protocol which enables the production of phage cocktails shown to be highly effective *in vivo*. Although this method produces highly effective phage cocktails, it does not solve the problems associated with intracellular pathogens or phage clearance by neutralizing antibodies during the course of treatment.

Two limitations of traditional phage therapy include the inability of phages to target intracellular bacteria and anti-phage antibodies which can neutralize and clear phages before they can reach their bacterial target. To address these limitations, researchers used liposome to entrap the phages, protecting them from antibodies and enabling the phages to be delivered to macrophages containing intracellular bacteria[138]. The results are promising, as the liposome encapsulated phages were 100 % resistant to phage-neutralizing antibodies and the liposomes were able to effectively deliver phages to the macrophages, resulting in 94.6 % reduction in intracellular *Klebsiella pneumoniae* concentrations. This research has addressed two problems associated with phage pharmacokinetics by enabling phages to access intracellular pathogens while also protecting them from neutralizing antibodies.

These fundamental studies have set the groundwork for scientists and companies to formulate better phage cocktails which stave off receptor mutations, target intracellular pathogens and have longer half-lives *in vivo*.

Phage therapy clinical trials

Although phage therapy has been successfully used in parts of Europe since before the dawn of antibiotics[139], more research must be completed to have it approved as a viable treatment option in North America. With a rise in antibiotic resistance globally, this research is more pertinent to our society than ever before. Recently the World Health Organizations' Antimicrobial Resistance Global Report on Surveillance stated a post-antibiotic era is becoming a real possibility for the 21st century due to the rise in antibiotic resistance[10]. There are nine clinical trials registered with the Food and Drug Administration (FDA) which are summarized in Table 1-1. Three clinical trials were not registered with the FDA. In total, 12 clinical trials have been conducted and each one is detailed below.

Three important clinical trials were published in 2009 which showed phages are safe and effective treatment options for clinical use. Merabishvili et al.[140] developed a phage cocktail which contained lytic phages against *P. aeruginosa* (podovirus PNM and myovirus 14/1) and *Staphylococcus aureus* (myovirus ISP). The cocktail was used on eight patients at a burn center in Brussels, Belgium and the applications were found to be safe, though no additional details on the efficacy were described. During that same year, results from a Phase I trial were published which focused on the treatment of chronic venous leg ulcers at a wound center in Lubbock, Texas (NCT00663091)[141]. The study tested a phage cocktail containing eight lytic myoviruses and podoviruses specific for *P. aeruginosa*, *Staphylococcus aureus*, and *Escherichia coli* on 18 patients using ~ 10^9 plaque forming units (PFU)/ml. The study found phage administration to be safe, though no significant difference was observed between the test and control groups for frequency of adverse events, rate of healing, or frequency of healing.

The first phage therapy Phase I/II clinical trial conducted in the United Kingdom to determine the safety and efficacy of phages[142]. The Biophage-PA phage cocktail, which consisted of six lytic phages against *P. aeruginosa,* was used to treat antibiotic-resistant *P. aeruginosa* infections of patients suffering from chronic otitis externa. The patients had experienced chronic otitis externa from 2 to 58 years. Prior to treatment, Biophage-PA was tested

against each patients' *P. aeruginosa* strain to ensure at least one of the phages from the cocktail could successfully infect it. The cocktail was delivered in a single 200 µl dose consisting of 6 x 10^5 plaque forming units (PFU). The single application completely cured three patients by the end of the trial, while all other patients saw improvement in their symptoms within six weeks. The researchers found statistically significant improvements in the phage-treated patients compared to the placebo group, and no treatment related adverse reactions occurred due to the phage application.

In 2011 a clinical trial (NCT01818206) began whereby researchers collected 58 sputum samples from cystic fibrosis (CF) patients for the purposes of studying phage susceptibility of the *P. aeruginosa* strains present in the sputum[143]. A cocktail of ten bacteriophages was used against the isolated strains to determine potential efficacy of the phages in sputum. Researchers found 45.8 % of the samples showed a decrease in cell numbers which was associated with an increase in bacteriophage concentrations. Additionally, each of the ten phages was assessed against each strain individually, which revealed 64.6 % of the strains were susceptible to bacteriophage. The researchers viewed this outcome as a major step forward in the development of phage therapy for chronic lung infections.

Results from a phage therapy clinical trial (NCT00937274) were published in 2016 from Bangladesh which focused on the treatment of acute bacterial diarrhea in children cause by enterotoxigenic *Escherichia coli* (ETEC) and/or enteropathogenic *Escherichia coli* (EPEC)[144]. The researchers orally administered either a T4-like coliphage cocktail consisting of 11 phages (Ab2, 4, 6, 11, 46, 50, 55, JS34, 37, 98, D1.4), a commercial Russian coliphage product with 18 phages (Microgen ColiProteus phage cocktail), or a placebo. The phages or placebo was administered over 4 days while the safety tolerance, stool volume, and stool frequency were monitored. The study showed no adverse events due to oral phage application, but the phages failed to amplify in the intestines and did not improve the diarrhea outcome, possibly due to insufficient phage coverage or too low of phage dosage.

The next Phase I/II clinical trial (NCT02116010) using the product PHAGOBURN (Pherecydes Pharma) started July 2015 and ended July 2016, though no research has been published on the outcomes of this trial. The goal of the study was to determine the tolerance and efficacy of a bacteriophage treatment on burn wound infections caused by *E. coli* or *P. aeruginosa*. No details about the phages used were provided. The research was funded by the

European Commission and involved seven clinical sites in the European Union. An additional study out of Europe conducted at the Polish Academy of Sciences located in Wroclaw, Poland performed a clinical trial (NCT00945087) using a group of phages against non-healing postoperative wounds or bone, upper respiratory tract, genital or urinary tract infections in which extensive antibiotic therapy failed or the use of the targeted drug is contraindicated. The start date and end date of the study were not provided, and no additional details on the types of phages used or the bacteria targeted was given.

A collaboration between the PhageBiotics Research Foundation and two American hospitals located in Washington state resulted in a publication on compassionate-use phage therapy of nine American patients with diabetic toe ulcers[145]. After failure of antibiotic therapy in every patient, their only option for treatment was amputation. Instead of amputation, the nine patients were approved for compassionate care use of phages and were administered the topical *Staphylococcus aureus* myovirus Sb-1. All patients healed over approximately seven weeks, except one patient who required 18 weeks. The phage treatment prevented amputation in all nine patients and was a promising step forward for phage therapy in the West as it illustrated the efficacy of phage treatment in hard to treat wounds. To further build on the diabetic toe ulcer study, a French Phase I/II clinical trial (NCT02664740) is planned to start in January 2019 and finish August 2019. The trial will use PhagoPied (Pherecydes Pharma), a gauze impregnated with bacteriophages against *S. aureus.* The study will enroll 60 patients with diabetic foot ulcers to compare the efficacies of the phage treatment versus a sterile compress. The PhagoPied gauze, containing a phage solution at 10^7 PFU/mL, will be applied on days 0, 7, and 14, unless the wound is already healed.

In 2017, Leitner et al.[146] published a double-blind clinical trial (NCT03140085) proposal for the investigation of bacteriophages against urinary tract infections of patients undergoing transurethral resection of the prostate. Patients in the trial will be randomized and receive either a bacteriophage solution, placebo, or antibiotic treatment over seven days. The authors noted that over the seven days, the bacteriophage cocktail will be subjected to periodic adaptation cycles to optimize the treatment with phages. This was not conducted with other clinical trials and may produce more favourable outcomes for each patient than using the standard cocktail alone.

Finally, the last two clinical trials to discuss from Table 1-1 are trials which are being conducted by AmpliPhi Biosciences Corporation and the United States Army Medical Research

and Material Command. These trials are an FDA Expanded Access regulation and only patients under the care of a licensed physician with expertise and facilities appropriate for the administration of the investigational medicine will be considered. The Expanded Access regulation was granted for serious or immediate life-threatening *S. aureus* (NCT03395769) and *P. aeruginosa* (NCT03395743) infections. No additional information was provided regarding the types of phages used other than the name of the cocktails; AB-SA01 and AB-PA01 for *S. aureus* and *P. aeruginosa* infections respectively. This is the first FDA Expanded Access trial granted for phage therapy.

In summation, there have been 12 clinical trials completed focusing on the clinical application of phages to treat antibiotic-resistant infections. The results from many of the trials have been resoundingly positive, apart from the EPEC/ETEC clinical trial performed in Bangladesh. The paper detailing the Bangladesh clinical trial highlighted potential reasons for the outcomes of the trial such as poor phage coverage and too low of phage titer. This suggests optimizing the phage cocktails over the course of treatment may be beneficial in future clinical trials. Recently, the first phage therapy center was opened in North America called the Center for Innovative Phage Applications and Therapeutics (IPATH), located at the University College San Diego School of Medicine. The opening of IPATH demonstrates the support behind phage therapy in the West, and it comes just in time to help combat the antibiotic resistance crisis our world is currently facing.

Table 1-1: Clinical trials in phage therapy.

Trial number	Study title	Status	Trial summary
NCT02664740	Standard treatment associated with phage therapy versus placebo for diabetic foot ulcers infected by *S. aureus* (PhagoPied)	Not yet recruiting	• Multicenter trial comparing phage impregnated dressing (10^7 PFU/mL) to a placebo dressing. • Dressings replaced at Day 7 and 14. • Wound healing measurements at week 12 • Monitor presence/absence of bacteria and antibiotic resistance

NCT00945087	Experimental phage therapy of bacterial infections	Unknown status	• Unknown bacterial targets • Non-healing wounds, upper respiratory tract infections, genital or urinary tract infections which failed antibiotic treatment.
NCT02116010	Evaluation of phage therapy for the treatment of *Escherichia coli* and *Pseudomonas aeruginosa* wound infections in burned patients (PHAGOBURN)	Unknown status	• Phase I/II multicenter trial comparing phage cocktails against silver sulfadiazine. • Time taken to get persistent reduction of bacteria compared to abundance at Day 0. • Assess tolerance to phage and level of clinical improvement.
NCT01818206	Bacteriophage effects on *Pseudomonas aeruginosa* (MUCOPHAGES)	Completed[143]	• Induced sputum samples from 59 CF patients. • *P. aeruginosa* count at 6 and 24 h post-exposure. • Phage counts after 6 h.
NCT00937274	Antibacterial treatment against diarrhea in oral rehydration solution	Terminated[144]	• Comparison of two T4 phage cocktails against oral rehydration solutions in ETEC and EPEC infections. • Assessment included safety tolerance and reduction of stool volume and frequency.
NCT03140085	Bacteriophages for treating urinary tract infections in patients undergoing transurethral resection of the prostate	Recruiting	• Phase 2|Phase 3 • Periodic adaptation cycles to optimize phage treatment. • Unknown bacterial targets.
NCT03395769	Individual patient expanded access for AB-SA01, an investigational anti-*Staphylococcus aureus* bacteriophage therapeutic	Available	• Expanded Access program

NCT03395743	Individual patient expanded access for AB-PA01, an investigational anti-*Pseudomonas aeruginosa* bacteriophage therapeutic	Available	• Expanded Access program
NCT00663091	A prospective, randomized, double-blind controlled study of WPP-201 for the safety and efficacy of treatment of venous leg ulcers	Completed[141]	• Phase I safety study evaluating therapeutic cocktail of eight phages (~ 10^9 PFU/mL/phage).

Characterization of phage for therapy

In most of the clinical trials completed to date, each phage included in the phage cocktail was fully sequenced and characterized prior to their use in therapy. Proper characterization of the phages before use is of utmost importance to ensure patient safety. Characterization of phages involves identifying their lifestyle, sequencing and annotating their genome, and determining their host range and the receptor used. Genome sequencing is extremely important to identify any genes involved in virulence or antibiotic resistance as they will be released into the environment upon host cell lysis. Determining phage lifestyle is also important as temperate phages have been shown to play a significant role in bacterial pathogenicity. For example, identification of the CTXΦ filamentous phage, which encodes the cholera toxin of *Vibrio cholerae*, highlighted the role of temperate phages in the evolution of bacterial pathogenicity[147].

Further, with the significant decrease in sequencing costs in the 1990s, sequenced bacterial genomes enabled researchers to identify prophages and phage-like elements in the majority of bacterial pathogens such as *V. cholerae*[147], *Streptococcus pyogenes*[148], and *Escherichia coli* O157:H7 str. Sakai[149]. Approximately 10 % of the *S. pyogenes* genomes consist of prophages which encode multiple virulence factors, and in the case of *E. coli* O157:H7 str. Sakai, 18 prophages were identified which accounts for 16 % of its genome. Ten randomly chosen *S. maltophilia* complete genomes from the NCBI database were analyzed with PHASTER[150], and the number of complete, questionable, and incomplete prophages identified ranged from zero (strain AA1) to eight (strain FDAARGOS_325) (data not shown). Researchers have noted that prophages account for the main genetic variability observed between closely

related bacterial strains[151]. This is demonstrated with the *S. maltophilia* prophage findings. Although bacteriophages can play a role in bacterial pathogenicity, choosing carefully screened phages for therapeutic use will ensure patient safety.

***Stenotrophomonas* phages**

The earliest research into bacteriophages against *S. maltophilia* began in 1973 by A. M. Moillo[152] who isolated and characterized the *Siphoviridae* phage M6 from a lysogenic strain of *Pseudomonas maltophilia.* The researcher wanted to isolate a transducing phage against *P. maltophilia* to compare transduction and gene arrangements against *P. aeruginosa* strains. While working with M6, Moillo passaged high-titer lysate on the *P. aeruginosa* strain PAO1. A phage mutant was isolated that could successfully infect PAO1; though, the mutant phage could no longer plaque on *P. maltophilia*. The mutant, named M6a, was shown to be capable of generalized transduction in *P. aeruginosa* strains, but due to the issues identified with a PAO1 grown M6a stock, transduction experiments between *P. aeruginosa* and *P. maltophilia* could not be completed. Following this article, nothing was published on *S. maltophilia* bacteriophages until 2005, at which time the purpose of the phage characterizations was for their use in therapy instead of as a genetic tool[153].

The second study published on *S. maltophilia* phages identified eight phages that were isolated from samples such as patient specimens, catheter-related devices, and wastewater[153]. One phage, ΦSMA5, was chosen for characterization. Electron microscopy of ΦSMA5 revealed a *Myoviridae* phage, with an isometric head, contractile tail, baseplate, and short tail fibers. The phage has a large genome estimated to be 250 kb and readily digestible with most restriction enzymes. ΦSMA5 is quite active against many *S. maltophilia* strains, successfully infecting 61 out of 87 strains tested. This phage may have been sequenced but it was not submitted to NCBI. The researchers highlighted ΦSMA5 shows potential for inclusion in a therapeutic phage cocktail.

The next phage documented was by Hagemann et al. (2006)[154] who published on their discovery of a self-replicating DNA molecule 6,907 bp in size isolated from chromosome preparations of a clinical *S. maltophilia* strain[154]. The scientists identified the DNA was the filamentous phage ΦSMA9, which was found to encode a zonula occludens-like toxin (zot) similar to the Zot toxin of *Vibrio cholerae*.

The next phage characterized was identified as a novel T4-like virulent phage isolated from a Kaohsiung, Taiwan sewage sample[155]. Characterization of the phage Smp14 revealed a 160 kb genome which encodes structural proteins 15-45% identical to structural proteins of the T4-like phages. Electron microscopy confirmed its *Myoviridae* morphology. The researchers submitted a 16 kb DNA sequence to NCBI. It is unknown why the entire genomic sequence was not submitted to NCBI for this phage.

Following the characterization of Smp14, Garcia et al. (2008) documented the isolation of 22 phages from sewage enrichment and prophage induction experiments[156]. Three of the phages were chosen for further analysis due to their host ranges. The S1 and S4 phages were isolated from prophage induction and belong to the *Siphoviridae* family. The S3 phage, which was isolated from a sewage sample, was shown to be an obligately lytic *Myoviridae* phage, and thus shows therapeutic potential. Only the S1 phages genomic sequence was submitted to NCBI.

After a four-year lull in publications, two groups of researchers from China published their findings on the novel *S. maltophilia* phages in 2012[157-159]. Two papers were published from the Beijing Institute of Microbiology and Epidemiology (BIME) group on two virulent phages named IME13 and IME15[157,158]. Both papers failed to report any electron microscopy work or host range studies; thus, the morphology for each phage can only be speculated. The IME13 phage features a large burst size (>3,000 phage/cell) and a unique plaque polymorphism with three plaque sizes[157]. The IME15 phage is T7-like due to the proteins identified from the genome analysis[158].

The second group to publish in 2012 detailed the novel phage φSHP1 isolated from the *S. maltophilia* strain P2[159]. Electron microscopy of φSHP1 showed a filamentous phage 2.1 μm long belonging to the *Inovirus* genus. The phage features a 6,867 bp long single-stranded DNA genome which also encodes a Zot-like toxin.

The following year, a translated abstract was published from a 2013 Chinese journal article documenting the characteristics of the *Myoviridae* phage SM1[160]. The genome was not sequenced, but gel electrophoresis suggests the genome is around 50 kb. Lysis experiments show a burst period of 50 min with an average burst of 187 phages per cell. The abstract also detailed the first *in vivo* mouse experiments involving *S. maltophilia* and phage. The results were promising with 100% of the SM1 treated mice surviving *S. maltophilia* infection past day 7.

In 2014, two papers were published on three temperate phages isolated from *S. maltophilia* lysogens. The phages ΦSMA6 and ΦSMA7 were identified as two novel filamentous phages of the *S. maltophilia* environmental strain Khak84[161]. Genomic analysis of the revealed both phages belong to the *Inoviridae* family and share a similar layout of functional gene groups compared to previously identified filamentous phages. The phages encode only 11 predicted open reading frames each and both encode Zot-like toxins. Following this, researchers published on a temperate myophage Smp131 which was isolated from the clinical *S. maltophilia* lysogen T13[162]. Genome analysis classifies Smp131 as a P2-like phage due to the genomic organization, arrangement of operons, and amino acid similarity to P2. Host range analysis suggests Smp131 has a narrow tropism as it could only infect three strains; though, researchers failed to mention the number of strains screened.

In 2015, the Dennis lab published on two *Siphoviridae* phages, DLP1 and DLP2, which are detailed in Chapter 2 (Table 1-3)[163]. DLP1 and DLP2 have narrow tropism within *S. maltophilia* clinical isolates, infecting 8 and 9 of 27 strains tested (Table 1-2). These phages are closely related to *P. aeruginosa* phages vB_Pae-Kakheti_25, PA73, and vB_PaeS_SCH_Ab26; which suggested DLP1 and DLP2 may have extended host ranges. Like M6, both phages can infect across taxonomic orders, though in contrast to M6, DLP1 and DLP2 can propagate on their original *S. maltophilia* host following propagation on *P. aeruginosa* strains.

The Dennis lab published a second *S. maltophilia* paper in 2017, on the T4-like myovirus DLP6 (Chapter 6 and Table 1-3)[164]. DLP6 was found to infect 13 of 27 clinical isolates screened (Table 1-2). This phage was found to be distantly related to both the T4-like Enterobacteriaceae family of phages and the T4-like cyanobacteriophage family though genomic analysis of the core and universal core genes of the T4-like phage groups. Unlike other T4-like phages, DLP6 is predicted to encode a transposase, though no stable lysogens were recovered.

Also published in 2017 was the identification of phages isolated for their potential use in biotechnology to aid in rust-prevention. A biocorrosion causing strain of *S. maltophilia* (PBM-IAUF-2) was used to isolate two phages from silversmithing wastewater in Isfahan, Iran[165]. The two phages were not named, but they have *Siphoviridae* morphology with excessively long tails (> 600 nm) which can be seen with the transmission electron micrographs provided.

The last *S. maltophilia* phage characterized at the time of this review is the *Siphoviridae* phage DLP5, which is the fourth phage characterized against *Stenotrophomonas* by the lab

(Chapter 5 and Table 1-3) [166]. DLP5 is found to share very little identity (2 %) with any phages in the NCBI database; thus, characterization of DLP5 led to the establishment of the new genus *Delepquintavirus.* DLP5 is a temperate phage and is closely related to a phage which has not yet been published on, DLP3 (Chapter 3, Table 1-2, and Table 1-3). Both phages are capable of causing lysogenic conversion of their hosts. The final phage which has not been published on is DLP4 (Chapter 4, Table 1-2, and Table 1-3) a temperate *Siphoviridae* phage also capable of lysogenic conversion. What is interesting to note about DLP4 is even though a stable lysogenic cycle is possible, no obvious temperate genes were identified with the use of BLASTp, pfam, and CD-Search.

In summary, 20 phages have been isolated and characterized against *S. maltophilia*, with three of the phages isolated for their potential use in biotechnology and genetics. Nine phages (Sm1, IME13, IME15, S3, Sm14, and ΦSMA5, DLP1, DLP2, and DLP6) were isolated and characterized for their potential use in phage therapy. In fact, the SM1 phage was the first phage used in a murine model where it was found to provide 100 % protection against *S. maltophilia* infection. The virulent phages characterized to date for use in therapeutic phage cocktails do provide hope in finding an alternative treatment option against *S. maltophilia* infections. It is important to note that out of 20 phages isolated, 11 were identified as temperate. Some of the papers did not discuss testing the lifestyle of the phages, so this number could potentially be higher. Of the temperate phages characterized to date, all four filamentous phages were found to encode the Zot-like toxin common to *Vibrio cholerae*, and all of the temperate phages we have characterised against *S. maltophilia* have been shown to cause lysogenic conversion of their hosts. Thus, although phage therapy against *S. maltophilia* infections is a possibility, researchers must be rigorous in identifying each phages lifestyle to ensure temperate phages are not used in therapy due to these findings.

Summary

The Smc is a group of opportunistic pathogens which cause serious mortality in immunocompromised patients. Due to their innate antibiotic resistance, treatment of these infections is difficult and new therapeutic options must be explored. Bacteriophage therapy is a promising alternative to antibiotic use and has already shown to be active against *S. maltophilia in vivo*[160]. The development of safe phage cocktails to treat these infections begins with the

isolation and characterization of phages against the pathogen. Characterization of the phages though genome sequencing, lifestyle determination, and host range are vital first steps in the construction of safe and effective cocktails. The following chapters, summarized in Table 1-2 and Table 1-3 will detail the genomic characterization of three novel candidates for Smc phage therapy, while also discussing three temperate phages found to cause lysogenic conversion of their host. The chapters highlight the importance of fully characterizing phages for therapy.

Table 1-2: Host range analysis results for the *Stenotrophomonas* bacteriophages isolated from Edmonton soil samples.

***S. maltophilia* strains**	**Bacteriophages**					
	DLP1	**DLP2**	**DLP3**	**DLP4**	**DLP5**	**DLP6**
101	+	+	++	++	-	-
102	-	+	+++	++	+	+
103	+	+	++	+++	-	+
152	-	-	-	-	-	-
155	-	-	++++	++++	-	+++
174	-	-	++	-	-	-
176	-	-	++++	-	-	+
213	+	+	++	-	-	+++
214	-	-	++	-	-	-
217	-	-	-	-	-	++
218	+	-	+	-	-	-
219	-	-	+++	++	-	+
230	-	-	++++	+	-	+
236	+	+	+	-	-	-
242	+	+	++	-	-	-
249	+	+	+	-	+++	-
278	-	-	-	+	-	-
280	-	++++	+++	++	-	-
282	-	-	+	++++	-	-
287	-	-	++	+	-	++
446	-	-	-	-	-	+
667	-	-	+	+	-	++
D1585	++++	++++	+++	+++	-	-
D1571	-	-	++++	-	++	+++
D1614	-	-	-	-	+++	-
D1576	-	-	+++	++++	++	++
D1568	-	-	-	+++	-	-
Total: 27	8	9	21	14	5	13

–, No sensitivity to phage; +, plaques at 10^{-2}; ++, clearing at 10^{-2}; +++, plaques at 10^{-4}; ++++, plaques at 10^{-6}

[a]Isolates from the Canadian *Burkholderia cepacia* complex Research Referral Repository

Table 1-3: Thesis overview of phages discussed in each chapter.

Chapter	Phage	Source, isolation strain	Length (bp)	GC (%)	Family	Relatedness	Unique features
2	DLP1	Red Deer River sediment, D1585	42,887	53.7	*Siphoviridae*	> 97 % to *P. aeruginosa* phages PA25, PA73, and Ab26[a]	• Host ranges cross taxonomic orders to *P. aeruginosa* strains. • Uses type IV pili as host receptor. • 57 ORFs DLP1, 58 ORFs DLP2
	DLP2	Blue flax soil, D1585	42,593	53.7	*Siphoviridae*		
3	DLP3	Empty soil, D1571	96,852	58.3	*Siphoviridae*	< 2 % to *Xylella* phage Sano	• Temperate phage, possible phagemid. • Second member of the *Delepquintavirus genus.* • Causes lysogenic conversion of D1571. • Encodes functional erythromycin resistance protein. • 148 ORFs; 5 tRNAs
4	DLP4	Planter soil, D1585	63,945	65.1	*Siphoviridae*	*P. aeruginosa* phages AAT-1, PaMx2836 and PaMx7436	• Temperate phage. • B2 morphotype *Siphoviridae.* • Causes lysogenic conversion of host. • Encodes functional trimethoprim resistance protein (DHFR) and virulence factor YbiA. • 82 ORFs; 1 tRNA
5	DLP5	Empty soil, D1614	96,542	58.4	*Siphoviridae*	< 2 % to *Xylella* phage Sano	• Type strain of *Delepquintavirus*. • Temperate phage maintained as phagemid. • Causes lysogenic conversion of D1614. • Encodes putative erythromycin resistance protein. • 149 ORFs; 5 tRNAs
6	DLP6	Planter soil, D1571	168,489	55.8	*Myoviridae*	*Sinorhizobium* phage phiM12	• Divergent T4-like virus with features from both the Enterobacteriaceae and cyanobacteria phage families. • Transposase encoded. • 241 ORFs; 30 tRNAs.

Chapter 2

The isolation and characterization of two *Stenotrophomonas maltophilia* bacteriophages capable of cross-taxonomic order infectivity

Portions of this chapter have been submitted as:

Peters, D. L., K. H., Lynch, P. Stothard and J. J. Dennis. 2015. The isolation and characterization of two *Stenotrophomonas maltophilia* bacteriophages capable of cross-taxonomic order infectivity. BMC Genomics. **16**:644.

McCutcheon, J. G., D. L., Peters and J. J. Dennis. 2018. Identification and characterization of the type IV pili as the cellular receptor of broad host range *Stenotrophomonas maltophilia* bacteriophages DLP1 and DLP2. Viruses. **10**:338.

OBJECTIVES

The objectives of this project were twofold: a) sequence and characterize the genomes of DLP1 (vB_SmaS-DLP-1) and DLP2 (vB_SmaS-DLP-2), and to determine shared characteristics of *Siphoviridae P. aeruginosa* phages which use the type IV pili for host infection.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Five *S. maltophilia* and eight *P. aeruginosa* strains were acquired from the Canadian *Burkholderia cepacia* complex Research and Referral Repository (Vancouver, BC). The *S. maltophilia* strains used for isolation of phages from soil samples were D1585, D1571, D1614, D1576 and D1568. An additional 22 *S. maltophilia* strains were gifted from the Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services, for host range analysis. All strains were grown aerobically overnight at 30 °C on half-strength Luria-Bertani (½ LB) solid medium or in ½ LB broth with shaking at 225 RPM.

Phage isolation, propagation, and electron microscopy

DLP1 and DLP2 were isolated from Red Deer River sediment and *Linum lewisii* (blue flax) soil, respectively, using standard extraction protocols[167]. Environmental samples were incubated with shaking at 30 °C in ½ LB broth, modified suspension medium (SM) (50 mM Tris–HCl [pH 7.5], 100 mM NaCl, 10 mM $MgSO_4$), and *S. maltophlia* D1585 liquid culture. Solids were pelleted by centrifugation and the supernatant was filter-sterilized using a Millex-HA 0.45 µm syringe driven filter unit (Millipore, Billerica, MA). These were plated in soft agar overlays with strain D1585 and incubated overnight at 30 °C followed by observation for >24 h at room temperature. For each environmental sample, a single plaque was isolated using a sterile Pasteur pipette, suspended in 500 µl of modified SM with 20 µl chloroform and incubated 1 h at room temperature to generate stocks for DLP1 and DLP2.

Propagation of the phages was performed using soft agar overlays: 100 µl liquid culture and 100 µl phage stock were incubated 20 min at room temperature, mixed with 3ml 0.7 % 1⁄2 LB top agar, overlaid on a plate of ½ LB solid medium, and incubated at 30 °C until plaque formation was complete. High titre stocks were made by overlaying plates exhibiting confluent

lysis with 3 ml modified SM and incubated >1 h at room temperature on a platform rocker. The supernatant was recovered, pelleted by centrifugation for 5 min at 10,000 × g, filter-sterilized using a Millex-HA 0.45 µm syringe-driven filter unit (Millipore, Billerica, MA) and stored at 4 °C. Titre of stocks was obtained using serial dilutions of phage stock into SM, followed by soft agar overlay technique and incubation at 30 °C until plaque formation was complete.

For electron microscopy, phage stocks were prepared as described above with the following modifications: ½ LB agarose plates and ½ LB soft agarose were used for overlays, MilliQ-filtered water for phage recovery and a 0.22 µm filter was used for syringe-driven filtration. A carbon-coated copper grid was incubated with lysate for 2 min and stained with 4 % uranyl acetate for 30 s. Transmission electron micrographs were captured using a Philips/FEI (Morgagni) transmission electron microscope with charge-coupled device camera at 80 kV (University of Alberta Department of Biological Sciences Advanced Microscopy Facility). The capsid diameter and tail length were measured with ImageJ and results were calculated using Microsoft Excel based on measurements from nine individual virions.

Host range analysis and PCR confirmation

Host range analysis was performed using a panel of 27 clinical *S. maltophilia* and 19 *P. aeruginosa* strains (Table 2-1 and Table 2-2), and 25 other Gram-negative bacterial species. Soft-agar overlays containing 100 µl liquid bacterial culture were allowed to solidify for 10 min at room temperature. These plates were spotted with 10 µl drops of DLP1 or DLP2 at multiple dilutions and assayed for clearing (confluent phage lysis), and/or plaque formation after incubation for 36 h at 30 °C. If plaques were formed, a single plaque from a successful infection plate was picked to propagate as a working stock solution for further analysis. To confirm the plaque contained DLP1 or DLP2 particles, PCR analysis was conducted on each purified phage solution using TopTaq DNA polymerase and buffers (Qiagen) and primers specific to each phage (DLP1F: ACACTGGCGAAGGATTACGG, DLP1R: GCCTTTCGAAATTCGCCGTT and DLP2F: CGGCTTTTTCGTGCCTGTAA, DLP2R: ACTC CTTTTCGATGCGTCCG) (Sigma-Genosys, Oakville, ON). These PCR products correspond to regions of DNA encoding part of ORF38, ORF39 and ORF40 in DLP1 and part of ORF38 and ORF39 in DLP2. PCR products were separated and visualized on a 1 % (wt/vol) agarose gel in 1x TAE (pH 8.0), and the product authenticity was confirmed by DNA sequencing. This test is a control experiment

designed to ensure that the application of an exogenous phage does not induce a resident prophage into production. All samples that were positive for the production of phage clearing or plaques were subjected to confirmation of DLP1 or DLP2 phage production by PCR.

Phage DNA isolation, RFLP analysis, and sequencing

DLP1 and DLP2 genomic DNA was isolated from bacteriophage lysate using the Wizard Lambda DNA purification system (Promega Corp., Madison, WI) with a modified protocol [168,169]. A 10 ml aliquot of filter-sterilized phage lysate (propagated on D1585 with agarose medium) was treated with 10 µl DNase I (Thermo Scientific, Waltham, MA), 100 µl 100x DNase I buffer (1 M Tris–HCl, 0.25 M $MgCl_2$, 10 mM $CaCl_2$), and 6 µl RNase (Thermo Scientific) and incubated 1 h at 37 °C to degrade the contaminating bacterial nucleic acids. Following incubation, 400 µl of 0.5 M EDTA and 25 µl of 20 mg/ml proteinase K (Applied Biosystems, Carlsbad, CA) were added and incubated 1 h at 55 °C to inactivate DNase I. After cooling to room temperature, the lysate was added to 8.4 g of guanidine thiocyanate, along with 1 ml of 37 °C resuspended Wizard DNA Clean-Up Resin (Promega Corporation, Madison, WI). This mixture was rocked for 10 min, and then pelleted by centrifugation at room temperature for 10 min at 5000 x g. The supernatant was drawn off until ~5 ml remained. This mixture was resuspended by swirling, transferred into a syringe attached to a Wizard Minicolumn (Promega Corporation), and pushed though the column. The column was then washed with 2 ml 80 % isopropanol and dried by centrifugation for 2 min at 10,000 x g. Phage DNA was eluted from the column following a 1 min incubation of 100 µl of 80 °C nuclease-free water (Integrated DNA Technologies, Coralville, IA) and centrifugation for 1 min at 10,000 x g. A NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA) was used to determine purity and concentration of eluted DNA.

For each phage DNA sample, restriction fragment length polymorphism analysis (RFLP) was performed using three - 20 µl FastDigest EcoRI (Thermo Scientific) reactions containing l µg of phage DNA. Reactions were incubated at 37 °C for 5 min and separated on a 1 % (wt/vol) agarose gel in 1x TAE (pH 8.0). Preliminary sequencing of EcoRI phage DNA fragments cloned into pUC19 was performed as described previously [170,171]. Phage DNA was submitted to The Applied Genomics Core at the University of Alberta for sequencing using MiSeq (Illumina, San Diego, CA) and assembled using the CLC Genomics Workbench (Qiagen, Toronto, ON). The

genome sequences of DLP1 and DLP2 have been deposited in GenBank with the accession numbers KR537872 and KR537871, respectively.

Bioinformatic analysis

Open reading frames (ORFs) for each contig were identified using the GLIMMER plugin[172] for Geneious [173] using the Bacteria and Archaea setting, as well as GeneMarkS for phage [174]. Conserved domain searches were performed using CD-Search[175]. The contigs were annotated using BLASTn and BLASTp (for full genomes and individual proteins, respectively) [176]. BLASTx and PHAST were used to search for similar sequences in the GenBank database. Sequence comparisons were visualized using Circos [177] and NUCmer [178] with the following parameters: breaklen = 200, maxgap = 90, mincluster = 65, min- match = 20. Lysis protein analysis was performed using TMHMM for transmembrane region identification[179].

Experimentally determined pili-binding *Siphoviridae* phages were identified in a search of the literature and the corresponding genomic sequence was used to perform a conserved domain search (CD-Search) to identify the potential presence of a phage-tail_3 domain which is found within gp26 (central tail hub) of DLP1 and DLP2. The CD-Search database CDD v3.16–50369 PSSMs was used to identify phage-tail_3 domains above the expected E-value threshold of 0.01. Composition-based statistics adjustment was used. The identified protein featuring the phage-tail_3 domain for each phage was then used for a multiple sequence alignment to include DLP1 and DLP2 using the MUSCLE[180] plugin for Geneious. Two multiple sequence alignments were also performed with the top ten BLASTp results for gp26 of DLP1 and DLP2. For each MUSCLE alignment, the maximum number of iterations selected was 8, with the anchor optimization option selected. The trees from iterations 1 and 2 were not retained. The distance measure for iteration 1 was kmer6_6 and was pctid_kimura for subsequent iterations. The clustering method was UPGMB for all iterations.

RESULTS AND DISCUSSION

Isolation, host range, and morphology

Using *S. maltophilia* strain D1585, phages DLP1 and DLP2 were isolated from Red Deer River sediment and soil planted with blue flax (*Linum lewisii*), respectively. In 2015, the published *S. maltophilia* phages had been isolated from clinical settings, sewage samples and

lysogenic bacteria [153-159,162]; therefore, DLP1 and DLP2 were the first phages to be isolated from sediment and soil.

Phage DLP1 exhibits a unique plaque development that was previously identified in phages KL1 and AH2 that target bacteria of the *Burkholderia cepacia* complex[181]. As with KL1 and AH2, stocks of DLP1 can be concentrated (up to 10^{10} plaque forming units [PFU]/ml) but use of such high titre stocks results in plates with no plaques. Instead, when lower titres (10^{7} PFU/ml or less) are used, and the plates are incubated at 30 °C for at least 24 h, DLP1 plaque development occurs (Figure 2-1). Individual plaques for DLP1 are turbid with no distinct borders and a diameter of 0.4-1 mm, averaging 0.7 mm. This contrasts the plaque development of phage DLP2, which produces clearing at high titres and clearly defined plaques at lower titres following 16 h incubation at 30 °C (Figure 2-1). Plaque sizes for DLP2 are clear with distinct boarders and a diameter 0.2-0.8 mm, averaging 0.4 mm.

DLP1 and DLP2 are classified in the order *Caudovirales* and the family *Siphoviridae* due to their morphological characteristics observed using electron microscopy. The DLP1 phage has a long, non-contractile tail of ~ 175 nm in length and a capsid size of ~ 70 nm in diameter (Figure 2-2a). Phage DLP2 is larger, with a non-contractile tail of ~ 205 nm and a capsid size of ~ 70 nm in diameter (Figure 2-2b).

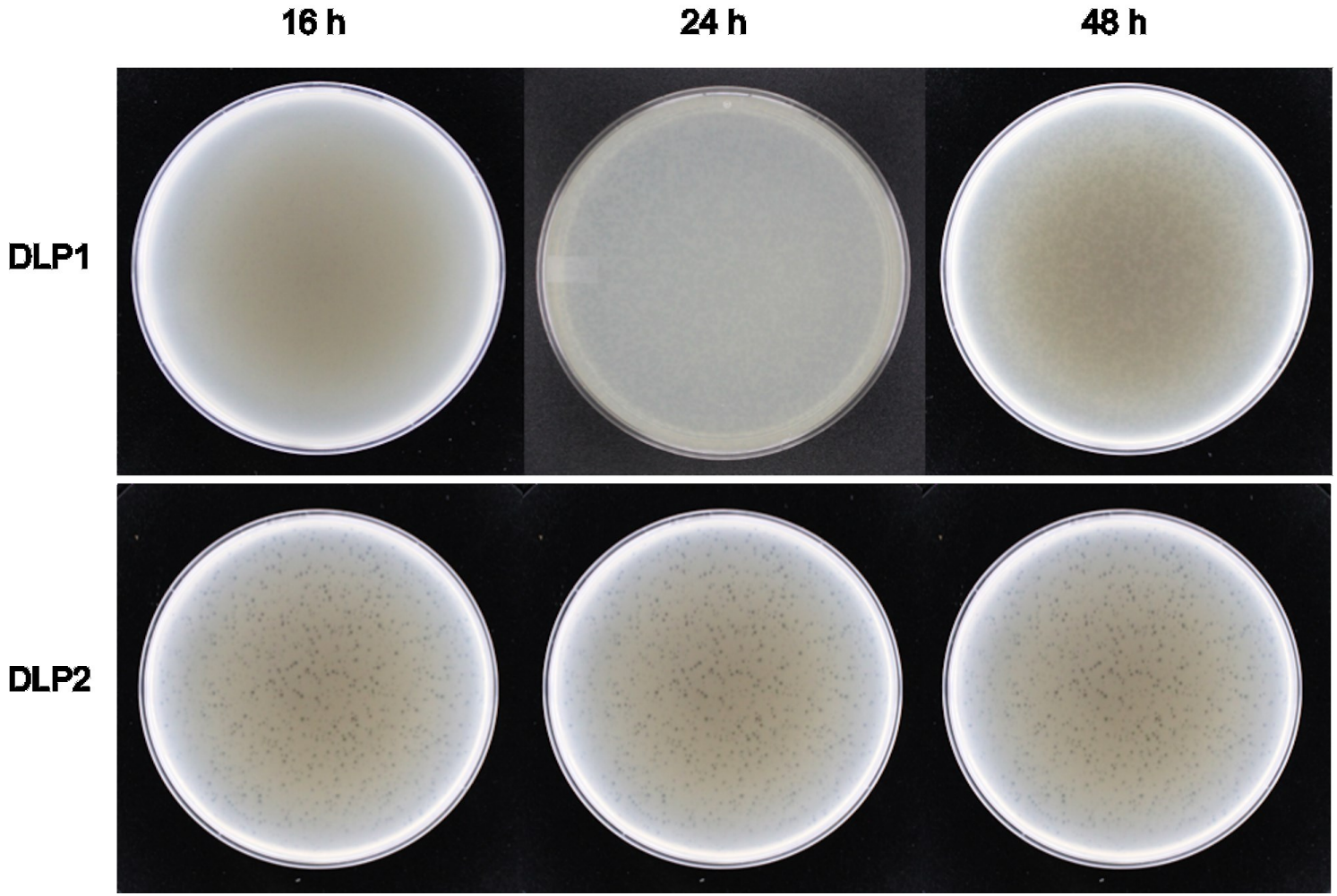

Figure 2-1: Development and morphology of DLP1 and DLP2 plaques. Phages were plated in half-strength Luria-Bertani (½ LB) agar overlays with 16 h liquid culture of *Stenotrophomonas maltophilia* D1585. Plates were incubated at 30 °C and photographed at 16, 24 and 48 h. Turbid DLP1 plaques were difficult to visualize until after 24 h of growth, whereas clear, well-defined DLP2 plaques were observed after 16 h.

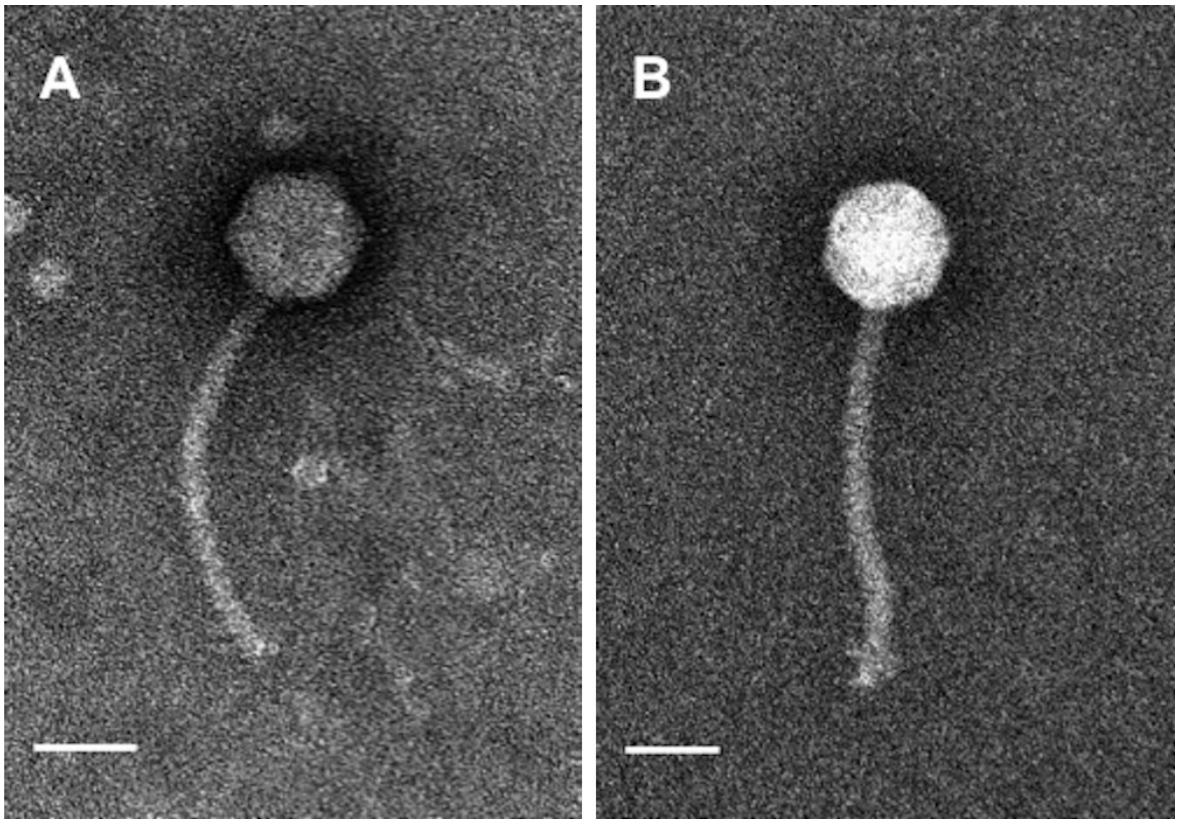

Figure 2-2: DLP1 (a) and DLP2 (b) phage morphology. Phages were stained with 4 % uranyl acetate and visualized at 180,000-fold magnification by transmission electron microscopy. Scale bars represent 50 nm. Both *Siphoviridae* family phages were of similar size, although the tail of DLP1 (175 nm) was shorter than that of DLP2 (205 nm).

Both DLP1 and DLP2 have a moderate host range within the *S. maltophilia* strains tested, with the ability to infect eight and nine out of 27 strains, respectively (Table 2-1). Both phages also have a unique ability to infect across bacterial taxonomic orders, with each phage capable of infecting two separate *P. aeruginosa* strains each (Table 2-2). The extended host range into *P. aeruginosa* strains is an interesting finding, as bacteriophages are typically thought to be relatively species specific. However, there are examples of bacteriophages that have been shown to lyse bacteria of different genera. For example, some phages originally discovered to infect one genus of cyanobacteria, have also been shown to be able to lyse other cyanobacteria genera [182-184]. A successful DLP1 and DLP2 infection and lysis of *P. aeruginosa* strains was confirmed with the use of PCR (Figure 2-3). DLP1 and DLP2 were also screened against an additional 25 Gram-negative bacteria which did not result in infection - *Pseudomonas fluorescens*: D1492, D1499, D1557, D1612; *Pseudomonas putida*: D0034, D1275, D1500, D1501; *Pseudomonas*

stutzeri: C9295, D0399, D0997, D1035; *Burkholderia multivorans*: ATCC17616, KLB; *Burkholderia cepacia*: C6433, C5393, K56-2; *Burkholderia gladioli*: CEP0029, CEP0071, CEP0082; and *Acinetobacter baumannii*: 17978, 19606, AYE, SDF, 1441-1.

Table 2-1: Host range analysis of DLP1 and DLP2 against *Stenotrophomonas maltophilia* clinical isolates.

S. maltophilia **strain**	**DLP1**	**DLP2**
101	+	+
102	–	+
103	+	+
152	–	–
155	–	–
174	–	–
176	–	–
213	+	+
214	–	–
217	–	–
218	+	–
219	–	–
230	–	–
236	+	+
242	+	+
249	+	+
278	–	–
280	–	++++
282	–	–
287	–	–
446	–	–
667	–	–
D1585[a]	++++	++++
D1571[a]	–	–
D1614[a]	–	–
D1576[a]	–	–
D1568[a]	–	–

–, No sensitivity to phage; +, plaques at 10^{-2}; ++, clearing at 10^{-2}; +++, plaques at 10^{-4}; ++++, plaques at 10^{-6}.
a Isolates from the Canadian *Burkholderia cepacia* complex Research Referral Repository

Table 2-2: Host range analysis of DLP1 and DLP2 against 19 *P. aeruginosa* strains.

***P. aeruginosa* strain**	**DLP1**	**DLP2**
PA01	++	–
HER1004	–	+++
HER1012	–	–
14715	–	++
Utah3	–	–
Utah4	–	–
14655	-	–
6106	-	–
pSHU-OTE	–	–
D1606D[a]	–	–
D1615C[a]	–	–
D1619M[a]	–	–
D1620E[a]	–	–
D1623C[a]	–	–
ENV003[a]	–	–
ENV009[a]	+++	–
FC0507[a]	–	–
R285	–	–
14672	–	–

–, No sensitivity to phage; +, plaques at 10^{-2}; ++, clearing at 10^{-2}; +++, plaques at 10^{-4}; ++++, plaques at 10^{-6}.
[a] Isolates from the Canadian *Burkholderia cepacia* complex Research Referral Repository.
[b] Isolates from cystic fibrosis patients.

Genome characterization

Genomic analysis of phages DLP1 and DLP2 reveals they are closely related phages. Initially, a comparison of restriction fragment length polymorphisms (RFLP) of DLP1 and DLP2 EcoRI-digested genomic DNA shows similar banding patterns with slight band differences between 0.85-1, 2-3 and 5 kbp (Figure 2-4). DLP1 and DLP2 similarity was confirmed by the results of whole genome sequencing using the Illumina platform (discussed below). A genome map for DLP1 and DLP2 (Figure 2-5) shows the modular layout of the two phages and their genetic similarity with respect to number of genes and genome size. However, complete genome sequencing also demonstrates the crudeness of RFLP analysis. The DLP1 genome contains 31 EcoRI sites, whereas the DLP2 genome possesses 32 EcoRI sites. Phage DLP1 possesses five DNA insertions of 29 bp in EcoRI fragment 6869-9910, 40 bp in fragment 9910-10,628, 50 bp in

fragment 12,987-13,729, 129 bp in fragment 24,500- 27,657 and 118 bp in fragment 34,709-39,879. Phage DLP2 possesses three DNA insertions of 40 bp in EcoRI fragment 10,559-11,971, 87 bp in fragment 14,003- 14,999, and 5 bp in fragment 29,984-31,617.

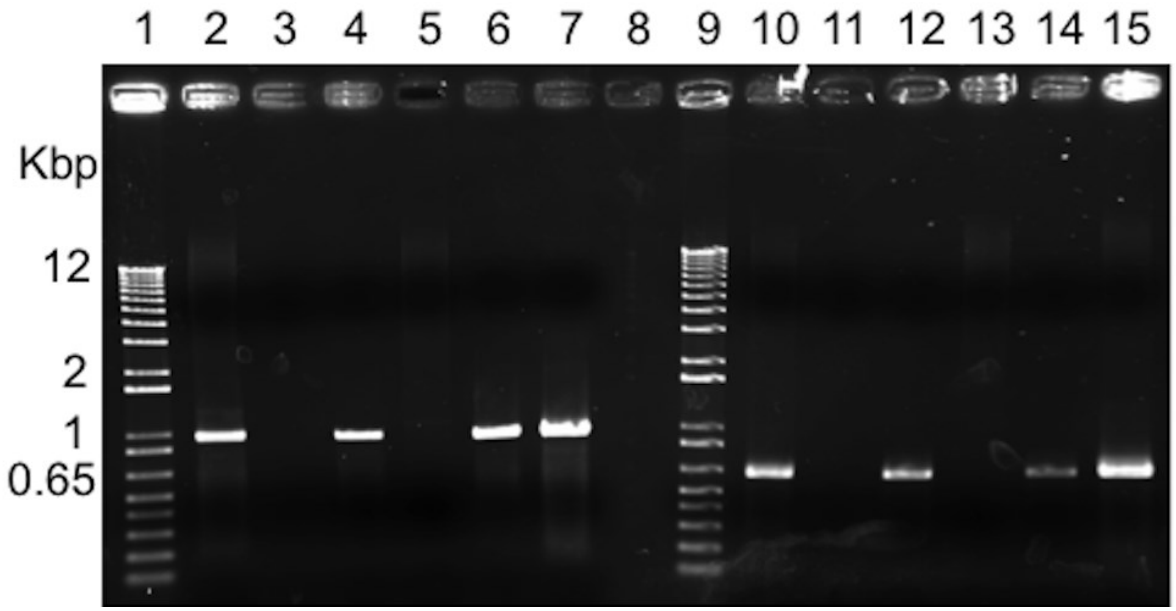

Figure 2-3: PCR confirmation of *P. aeruginosa* infections by DLP1 and DLP2. Lanes 1 and 9: 1 Kbp Plus DNA ladder (Invitrogen), lane 2: DLP1 DNA, lane 3: DLP1 negative control, lane 4: DLP1 lysate, lane 5: DLP2 DNA, lane 6: DLP1 lysate from PA01 infection, lane 7: DLP1 lysate from ENV009 infection, lane 8: blank, lane 10: DLP2 DNA, lane 11: DLP2 negative control, lane 12: DLP2 lysate, lane 13: DLP1 lysate, lane 14: DLP2 lysate from HER004 infection, lane 15: DLP2 lysate from 14715 infection. The size of the markers (in Kbp) is shown on the left.

In addition, phage DLP2 has an extra EcoRI site at base pair 3345 due to a point mutation. Phages DLP1 and DLP2 were found to be 96.7 % identical over 97.2 % of their genomes. However, this comparison still denotes considerable variation between the two phage genomic sequences. A BLASTn comparison indicates that the two genomes share 40,317 identical base pairs out of 41,687 aligned base pairs (1200 base pairs unaligned), with 166 gaps. The similarity of DLP1 and DLP2 to each other, and to their closest relative *Pseudomonas* phage vB_Pae-Kakheti25 (informally PA25), is illustrated in Figure 2-6, a Circos plot of a NUCmer comparison of the three phages.

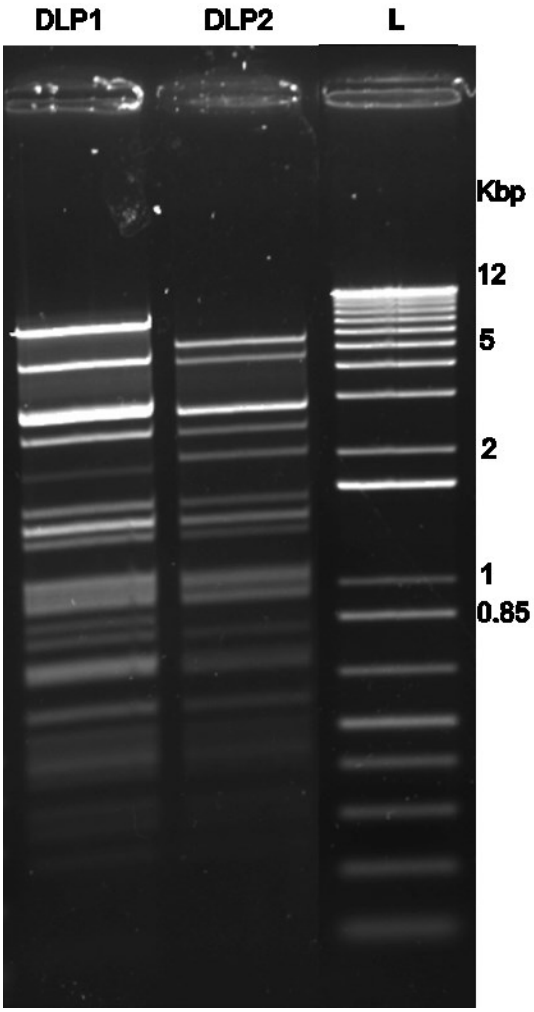

Figure 2-4: Restriction fragment length polymorphisms of DLP1 and DLP2 genomic DNA. 1 µg of phage DNA was digested 5 min with EcoRI and separated on a 1 % agarose gel. L: 1 Kbp Plus DNA Ladder (Invitrogen). Several banding pattern differences are apparent between the two phages.

The DLP1 genome is 42,887 base pairs (bp) in length, with a GC content of 53.7 %. DLP1 is predicted to encode 57 proteins with the most common start codon being ATG, although a TTG start codon is used for gp19, gp37 and gp41 (Table 2-3, Figure 2-5). Besides phage DLP2, phage DLP1 is most closely related to the siphovirus *Pseudomonas* phage vB_Pae-Kakheti25 (PA25) (Figure 2-6)[185]. DLP1 and PA25 phages are similar with respect to genome length (42,844 bp for PA25 and 42,887 bp for DLP1), GC content (53.7 % for PA25 and DLP1) and predicted number of proteins (58 for PA25 and 57 for DLP1)[185]. BLASTn comparison of DLP1 and PA25 shows a 98 % identity over 94 % of the genome. The genome of DLP1 also shows high similarity to phages PA73 (98 % identity over 92 % of the genome)[186] and

vB_PaeS_SCH_Ab26 (Ab26) (96 % identity over 92 % of the genome)[185]. Again, this represents a significant amount of genetic variation, with 145 gaps required to complete the genomic alignment with PA25, 144 gaps required to align PA73, and 220 gaps required to align Ab26, suggesting that although these phages belong to the same family, they are not merely variants of one another. The predicted proteins of DLP1 are similar to those found in phages PA25, PA73 and Ab26; though DLP1 proteins gp32, gp45 and gp48 are unique in that they are not similar to any proteins found in PA25, PA73 and Ab26 (Table 2-3). The DLP1 protein gp32 has identity to gp055 of the *Erwinia* phage vB_EamP-S6. The gp45 protein is related to the hypothetical protein X805_23910 of *Sphaerotilus natans* subsp. *natans* DSM 6575, which is a filamentous bacterium known to contribute to the stability of *Pseudomonas* sp. colonies at low concentrations[187]. The Vsr endonuclease encoded shares the highest identity to a protein found in the *Burkholderia* phage KL1[181].

The DLP2 genome is 42,593 bp in length, with a GC content of 53.7 %. DLP2 is predicted to encode 58 proteins with the most common start codon being ATG, although a TTG start codon is used for gp19 and gp37, and a GTG start codon is used for gp43 and gp55 (Table 2-4, Figure 2-5). Phage DLP2 is also related to *Pseudomonas* phage PA25 (Figure 2-6). These two phages are similar with respect to genome length (42,844 bp for PA25 and 42,593 for DLP2), GC content (53.7 % for PA25 and DLP2) and predicted number of proteins (58 for PA25 and DLP2) [179]. BLASTn comparison of the DLP2 and PA25 genomes shows a 97 % identity over 95 % of the genome. The BLASTn results also reveals DLP2 to be similar to *Pseudomonas* phages PA73 (98 % identity over 93 % of genome) [186] and Ab26 (97 % identity over 90 % of the genome)[185]. Phage DLP2 gene content differs from phages PA25, PA73 and Ab26 in predicted proteins gp39 and gp45 (Table 2-4). DLP2 protein gp39 is most closely related to the uncharacterized protein MAM_066 of the *Serratia* phage ΦMAM1. Similar to DLP1, the DLP2 gp45 protein is related to the hypothetical protein X805_23910 of *Sphaerotilus natans* subsp. *natans* DSM 6575.

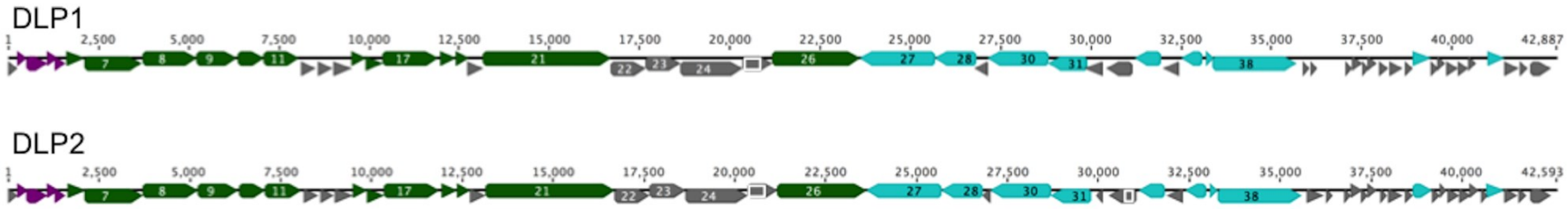

Figure 2-5: Genome maps of DLP1 and DLP2. The scale in kb is shown above the genome maps. The assigned functions for the predicted proteins encoded is as follows: grey, unknown function; purple, lysis; green, virion morphogenesis; blue, DNA replication/repair. Numbers within the larger ORFs are the gene product numbers.

Table 2-3: Bacteriophage DLP1 genome annotations.

Gene	Coding region	Putative function	Strand	Start codon	Length (AA)	Closest relative	Percent identity	Source	GenBank accession number
1	1-255	Hypothetical protein	+	ATG	84	ORF001	100	PA73	YP_001293408.1
2	252-518	Holin	+	ATG	88	Holin	99	PK25	YP_006299866.1
3	511-1056	Endolysin	+	ATG	181	Endolysin	98	PK25	YP_006299867.1
4	1068-1373	Rz	+	ATG	101	Rz	99	Ab26	YP_009044338.1
5	1288-1569	Rz1	+	ATG	93	Rz1	100	PK25	YP_006299869.1
6	1627-2115	Small terminase	+	ATG	162	Small terminase	99	PK25	YP_006299870.1
7	2096-3691	Large terminase	+	ATG	531	Large terminase	100	Ab26	YP_001293413.1
8	3705-5210	Portal protein	+	ATG	510	Portal protein	98	Ab26	YP_009044342.1
9	5222-6316	F-like head morphogenesis protein	+	ATG	364	ORF008	100	PA73	YP_001293415.1
10	6353-7072	Scaffold protein	+	ATG	239	Scaffold protein	100	PK25	YP_006299874.1
11	7075-8052	Major capsid protein	+	ATG	325	ORF010	99	PA73	YP_001293417.1
12	8122-8526	Hypothetical protein	+	ATG	135	ORF011	100	PA73	YP_001293418.1

13	8592-8993	Hypothetical protein	+	ATG	133	ORF12	71	Ab26	YP_009044347.1
14	9005-9523	Hypothetical protein	+	ATG	172	Hypothetical protein	92	Ab26	YP_009044348.1
15	9527-9907	Head-tail joining protein	+	ATG	126	Hypothetical protein	98	PK25	YP_006299879.1
16	9904-10359	Minor tail protein	+	ATG	151	ORF015	97	PA73	YP_001293422.1
17	10372-11907	Major tail tube protein	+	ATG	511	Major tail tube protein	99	PK25	YP_006299881.1
18	11971-12399	Tail chaperonin	+	ATG	142	ORF017	100	PA73	YP_001293424.1
19	12408-12764	Tail chaperonin	+	TTG	118	Tail chaperonin	100	PK25	YP_006299882.1
20	12733-13167	Hypothetical protein	+	ATG	144	ORF019	100	PA73	YP_001293426.1
21	13173-16700	Tape measure protein	+	ATG	1175	Tape measure protein	96	PK25	YP_006299885.1
22	16701-17663	Hypothetical protein	+	ATG	320	Hypothetical protein	87	PK25	YP_006299886.1
23	17663-18628	Hypothetical protein	+	ATG	321	Hypothetical protein	64	PK25	YP_006299887.1
24	18634-20346	Hypothetical protein	+	ATG	570	Hypothetical protein	96	PK25	YP_006299888.1
25	20346-21170	Hypothetical protein	+	ATG	274	Hypothetical protein	99	PK25	YP_006299889.1
26	21174-23615	Central tail hub	+	ATG	813	Central tail ub	99	PK25	YP_006299890.1
27	23616-25667	DNA polymerase	-	ATG	683	DNA polymerase	99	PK25	YP_006299891.1
28	25679-26821	Replicative clamp	-	ATG	380	Replicative clamp	99	PK25	YP_006299892.1
29	26805-27161	Hypothetical protein	-	ATG	118	ORF028	97	PA73	YP_001293435.1
30	27166-28821	DEAD box helicase	-	ATG	551	ORF029	100	PA73	YP_001293436.1

31	28814-29911	RecB exonuclease	-	ATG	365	ORF030	100	PA73	YP_001293437.1
32	29817-30344	Hypothetical protein	-	ATG	175	gp055	41	EamP-S6[a]	YP_007005791.1
33	30423-31169	Hypothetical protein	-	ATG	248	Member of the DUF669 phage protein family	99	PK25	YP_006299897.1
34	31228-31944	RecA	-	ATG	238	RecA	99	Ab26	YP_009044366.1
35	31999-32439	Hypothetical protein	-	ATG	147	ORF033	99	PA73	YP_001293440.1
36	32516-33073	MazG	-	ATG	185	MazG	89	PK25	YP_006299900.1
37	33193-33399	Transcriptional regulator	+	TTG	68	ORF035	100	PA73	YP_001293442.1
38	33389-35710	Replicative primase/helicase	+	ATG	773	Replicative primase/helicase	99	PK25	YP_006299902.1
39	35862-36062	Hypothetical protein	+	ATG	66	Hypothetical protein	99	Ab26	YP_009044372.1
40	36107-36256	Hypothetical protein	+	ATG	49	KAK25_00040	96	PK25	YP_006299904.1
41	37055-37234	Hypothetical protein	+	TTG	59	Hypothetical protein ORF0038	92	PA73	YP_001293445.1
42	37231-37527	Hypothetical protein	+	ATG	98	ORF0039	98	PA73	YP_001293446.1
43	37524-37703	Hypothetical protein	+	ATG	59	Hypothetical protein	90	Ab26	YP_009044375.1
44	37678-37920	Hypothetical protein	+	ATG	80	Hypothetical protein	96	Ab26	YP_009044376.1
45	38003-38224	Hypothetical protein	+	ATG	73	X805_23910	56	DSM 6575[b]	KDB52021.1
46	38271-38645	Hypothetical protein	+	ATG	124	KAK25_00046	99	PK25	YP_006299910.1
47	38706-38927	Hypothetical protein	+	ATG	73	KAK25_00047	97	PK25	YP_006299911.1
48	38924-39415	Vsr endonuclease	+	ATG	163	Vsr endonuclease	78	KL1[c]	YP_006560795.1

49	39403-39618	Hypothetical protein	+	ATG	71	Hypothetical protein	99	Ab26	YP_009044380.1
50	39615-39794	Hypothetical protein	+	ATG	59	KAK25_00050	95	PK25	YP_006299914.1
51	39855-40154	Hypothetical protein	+	ATG	99	ORF0045	81	PA73	YP_001293452.1
52	40171-40461	Hypothetical protein	+	ATG	96	ORF0046	99	PA73	YP_001293453.1
53	40454-40687	Hypothetical protein	+	ATG	77	KAK25_00053	100	PK25	YP_006299917.1
54	40983-41450	dCMP deaminase	+	ATG	155	dCMP deaminase	98	PK25	YP_006299919.1
55	41456-41839	Hypothetical protein	+	ATG	127	ORF0050	100	PA73	YP_001293457.1
56	41874-42083	Hypothetical protein	+	ATG	69	ORF0051	96	PA73	YP_001293458.1
57	42167-42739	Hypothetical protein	+	ATG	190	ORF0052	99	PA73	YP_001293459.1

a *Erwinia* phage vB_EamP-S6
b *Sphaerotilus natans* subsp. *natans* DSM 6575
c KL1 is *Burkholderia* phage KL1

Table 2-4: Bacteriophage DLP2 genome annotations.

Gene	Coding region	Putative Function	Strand	Start Codon	Length (AA)	Closest relative	Percent Identity	Source	GenBank accession number
1	1-255	Hypothetical protein	+	ATG	84	Phage protein found in lysis cassettes	96	PK25	YP_006299865.1
2	252-518	Holin	+	ATG	88	Holin	95	PK25	YP_006299866.1
3	511-1056	Endolysin	+	ATG	181	ORF003	100	PA73	YP_001293410.1
4	1068-1373	Rz	+	ATG	101	Rz	100	Ab26	YP_009044338.1
5	1288-1569	Rz1	+	ATG	93	Rz1	100	Ab26	YP_009044339.1
6	1627-2115	Small terminase	+	ATG	162	Small terminase	100	PK25	YP_006299870.1
7	2096-3691	Large terminase	+	ATG	531	Large terminase	99	Ab26	YP_009044341.1

8	3705-5210	Portal protein	+	ATG	501	Portal protein	99	Ab26	YP_009044342.1
9	5222-6316	F-like head morphogenesis protein	+	ATG	364	F-like head morphogenesis protein	99	Ab26	YP_009044343.1
10	6353-7072	Scaffold protein	+	ATG	239	Scaffold protein	98	PK25	YP_006299874.1
11	7075-8052	Major capsid protein	+	ATG	325	ORF010	99	PA73	YP_001293417.1
12	8122-8526	Hypothetical protein	+	ATG	134	ORF011	99	PA73	YP_001293418.1
13	8592-8963	Hypothetical protein	+	ATG	123	ORF012	98	PA73	YP_001293419.1
14	8976-9494	Hypothetical protein	+	ATG	172	Virion protein	100	PK25	YP_006299878.1
15	9498-9878	Head-tail joining protein	+	ATG	126	Virion protein	98	Ab26	YP_009044349.1
16	9875-10330	Minor tail protein	+	ATG	151	ORF015	99	PA73	YP_001293422.1
17	10343-11878	Major tail tube protein	+	ATG	512	Major tail tube protein	99	PK25	YP_006299881.1
18	11942-12370	Tail chaperonin	+	ATG	143	Tail chaperonin	99	Ab26	YP_009044352.1
19	12379-12735	Tail chaperonin	+	TTG	119	Tail chaperonin	99	Ab26	YP_009044353.1
20	12704-13138	Hypothetical protein	+	ATG	145	ORF019	100	PA73	YP_001293426.1
21	13143-16711	Tape measure protein	+	ATG	1189	Tape measure protein	99	PK25	YP_006299885.1
22	16709-17671	Hypothetical protein	+	ATG	321	Virion protein	88	PK25	YP_006299886.1
23	17671-18636	Hypothetical protein	+	ATG	322	Virion protein	64	PK25	YP_006299887.1
24	18642-20354	Hypothetical protein	+	ATG	571	Virion protein	96	PK25	YP_006299888.1
25	20354-21178	Hypothetical protein	+	ATG	275	Virion protein	99	PK25	YP_006299889.1

26	21182-23623	Central tail hub	+	ATG	814	Central tail hub	99	PK25	YP_006299890.1
27	23624-25675	DNA polymerase	-	ATG	684	DNA polymerase	99	PK25	YP_006299891.1
28	25687-26829	Replicative clamp	-	ATG	381	Replicative clamp	97	PK25	YP_006299892.1
29	26813-27040	Hypothetical Protein	-	ATG	76	KAK25_00029	100	PK25	YP_006299893.1
30	27045-28700	DEAD box helicase	-	ATG	552	DEAD box helicase	99	Ab26	YP_009044363.1
31	28693-29790	RecB exonuclease	-	ATG	366	RecB exonuclease	99	Ab26	YP_009044364.1
32	29959-30120	Hypothetical protein	-	ATG	54	KAK25_00032	100	PK25	YP_006299896.1
33	30302-31054	Hypothetical protein	-	ATG	251	Member of DUF669 phage protein family	99	PK25	YP_006299897.1
34	31113-31829	Rec A	-	ATG	239	RecA	100	Ab26	YP_009044366.1
35	31884-32324	Hypothetical protein	-	ATG	147	Hypothetical protein	99	Ab26	YP_009044367.1
36	32401-32958	MazG	-	ATG	186	MazG	98	Ab26	YP_009044368.1
37	33078-33284	Transcriptional regulator	+	TTG	69	Hypothetical protein	99	Ab26	YP_009044369.1
38	33274-35595	Replicative Primase/Helicase	+	ATG	774	Replicative primase/helicase	99	PK25	YP_006299902.1
39	35741-36193	Hypothetical protein	+	ATG	151	MAM_066	54	ΦMAM1[a]	YP_007349045.1
40	36281-36430	Hypothetical protein	+	ATG	50	KAK25_00040	69	PK25	YP_006299904.1
41	36764-36955	Hypothetical protein	+	ATG	64	ORF038	94	PA73	YP_001293445.1
42	36952-37248	Hypothetical protein	+	ATG	99	ORF039	100	PA73	YP_001293446.1
43	37269-37424	Hypothetical protein	+	GTG	52	Hypothetical protein	92	Ab26	YP_009044375.1

44	37399-37641	Hypothetical protein	+	ATG	81	ORF040	99	PA73	YP_001293447.1
45	37748-37945	Hypothetical protein	+	ATG	66	X805_23910	58	DSM 6575[b]	KDB52021.1
46	37993-38367	Hypothetical protein	+	ATG	125	KAK25_00046	97	PK25	YP_006299910.1
47	38428-38649	Hypothetical protein	+	ATG	74	KAK25_00047	99	PK25	YP_006299911.1
48	38646-39182	Vsr endonuclease	+	ATG	179	KAK25_00048	100	PK25	YP_006299912.1
49	39170-39385	Hypothetical protein	+	ATG	72	Hypothetical protein	100	Ab26	YP_009044380.1
50	39382-39561	Hypothetical protein	+	ATG	60	KAK25_00050	98	PK25	YP_006299914.1
51	39622-39927	Hypothetical protein	+	ATG	102	ORF045	100	PA73	YP_001293452.1
52	39944-40234	Hypothetical protein	+	ATG	97	ORF046	99	PA73	YP_001293453.1
53	40227-40460	Hypothetical protein	+	ATG	78	KAK25_00053	99	PK25	YP_006299917.1
54	40531-40698	Hypothetical protein	+	ATG	56	ORF048	100	PA73	YP_001293455.1
55	40689-41156	dCMP deaminase	+	GTG	156	dCMP deaminase	97	PK25	YP_009044383.1
56	41162-41545	Hypothetical protein	+	ATG	128	ORF050	100	PA73	YP_001293457.1
57	41580-41789	Hypothetical protein	+	ATG	70	ORF051	100	PA73	YP_001293458.1
58	41873-42445	Hypothetical protein	+	ATG	191	ORF052	99	PA73	YP_001293459.1

[a] *Serratia* phage ΦMAM1

[b] *Sphaerotilus natans* subsp. *natans* DSM 6575

Analysis of modules

The proteins identified in DLP1 and DLP2 can be classified into three general categories: lysis, virion morphogenesis (including DNA packing and capsid/tail morphogenesis) and DNA replication/repair. The ORFs of DLP1 and DLP2 are syntenic, and the predicted proteins are similar with only a few variations from each other (Table 2-3 and Table 2-4), yet these two phages exhibit two completely different plaque development characteristics (Figure 2-1). It is also of interest to note that no genes encoding known or putative virulence factors were discovered in the genomes of phages DLP1 and DLP2, or any other related phages in this family. This makes DLP1 and DLP2 great candidates for inclusion in a phage cocktail.

Lysis

Genes putatively encoding the lysis proteins holin, lysin, Rz, Rz1 and a hypothetical protein have been identified in DLP1. A BLASTp search of predicted protein gp1 shows that it is similar to a phage protein family found in lysis cassettes that was identified in phage PA25. A BLASTp search also showed gp2 to be a putative holin protein similar to those identified in PA25 and PA73. Analysis of this gp2 protein with TMHMM revealed it has two transmembrane domains; thus, gp2 is a predicted to be a class II holin[179]. Gene product 3 is nearly identical to the endolysin of PA25. Gp4 and gp5 proteins are similar to the Rz protein of Ab26 and Rz1 of PA25 respectively. The Rz protein is a class II inner membrane protein with an N-terminal domain and Rz1 is a proline-rich outer membrane lipoprotein [188]. The Rz/Rz1 proteins contribute to lysis by fusing to the inner and outer membranes following holin and endolysin activity to facilitate phage release [189]. The gp4 protein is predicted to contain a single N-terminal transmembrane domain, a characteristic of Rz proteins [179,190]. LipoP analysis of gp5 shows a signal peptidase II cleavage site between amino acids 20 and 21, resulting in a 73 amino acid protein with 7 proline residues (9.6 % proline) [191].

The lysis predicted proteins identified in DLP2 are similar to those also identified in phage DLP1. However, there are also some differences. The gp3 of DLP2 is most closely related to ORF003 of PA73 and also the endolysin protein in PA25. Like DLP1, gp4 of DLP2 shares the highest percent identity to the Rz protein of phage Ab26. Although DLP2 gp5 required manual annotation, BLASTp analysis showed it shares the highest identity to Rz1 of Ab26, rather than phage PA25. However, LipoP analysis revealed the identical signal peptidase II cleavage site as

for phage DLP1 gp5. Analysis of the lysis module for DLP1 and DLP2 did not reveal why phage DLP1 exhibits delayed plaque development when compared to phage DLP2. One hypothesis is that gp32 of DLP1, encoding a hypothetical protein not found in DLP2 (most closely related to gp055 of *Erwinia* phage vB_EamP-S6), contributes to the delayed plaque development of DLP1. However, until a definitive function for the DLP1 gp32-encoded protein can be established, this hypothesis remains unproven.

Virion morphogenesis

As discussed above, DLP1 is closely related to phage DLP2, and phages PA25, Ab26 and PA73, whose proteins are generally uncharacterized. BLASTp analysis of the 20 genes involved in virion morphogenesis in DLP1 identified 13 genes encoding proteins with putative functions: two involved in DNA packaging, four involved in capsid morphogenesis and seven involved in tail morphogenesis. The DNA packaging proteins gp6 (small terminase subunit) and gp7 (large terminase subunit) are similar to the small terminase subunit of PA25 and large terminase subunit of Ab26 respectively. Protein gp8 shares a 98 % identity to the portal protein of Ab26. Although gp9 shares 100 % identity to uncharacterized ORF008 of PA73, it has been assigned a putative function due to its high identity to the F-like head morphogenesis protein of Ab26 (Table 2-3). Gp10 shares 100 % identity to the scaffold protein of PA25. The gp11 is most similar to ORF010 of PA73, but its high identity to the major capsid protein of Ab26 has allowed a putative function to be assigned to this protein (Table 2-3). The seven proteins involved in tail morphogenesis are gp15 (head-tail joining protein), gp16 (minor tail protein), gp17 (major tail protein), gp18 (tail chaperonin), gp19 (tail chaperonin), gp21 (tape measure protein) and gp26 (central tail hub). Both gp15 and gp16 have high identity to a virion protein of Ab26 and gp15 of PA73 respectively. These gene products have been assigned putative functions due to their similarities to the head-tail joining protein (gp15) and minor tail protein (gp16) of *Burkholderia* phage KL1 (Table 2-3)[181]. Protein gp17 shares a 99 % identity to the major tail protein of PA25. Gp17 is 100 % identical to gp17 of PA73 but has been assigned the putative function of tail chaperonin due to its identity to Ab26 tail chaperonin. Like gp17, gp18 is predicted to be a tail chaperonin protein and has 100 % identity to the PA25 tail chaperonin protein. Both gp21 and gp26 are closely related to the tape measure protein and central tail hub of PA25, respectively.

Analysis of DLP2 with BLASTp revealed the virion morphogenesis proteins are nearly identical to those of DLP1, with only minor differences (Table 2-4).

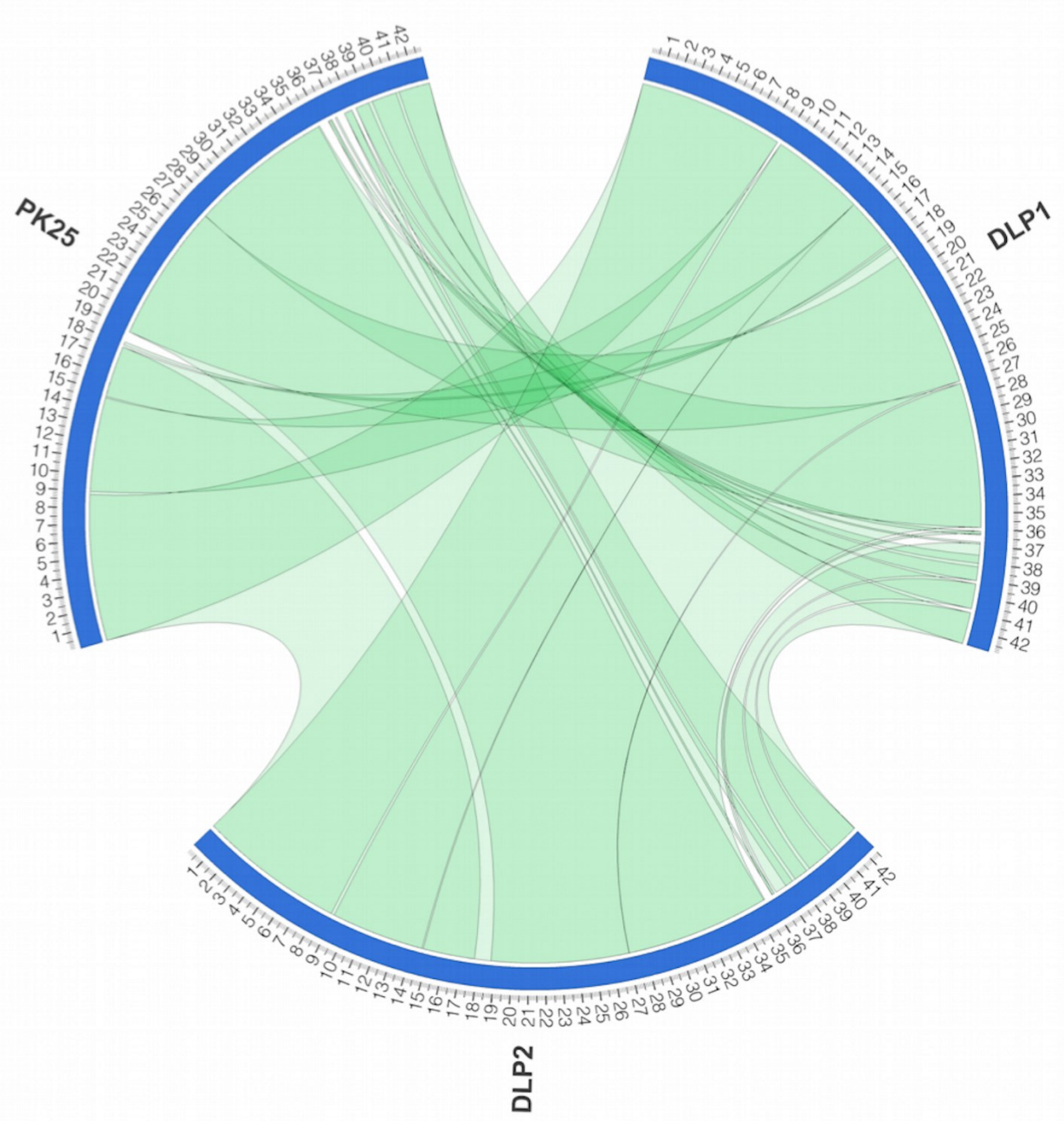

Figure 2-6: Circos plot of genomes from phages DLP1, DLP2, and vB_Pae-Kakheti25 (PA25) NUCmer comparisons. Green ribbon indicates regions of identity between the three genomes at the nucleotide level. The scale in kb is shown on the periphery of the plot for each phage. NUCmer parameters: breaklen = 200, maxgap = 90, mincluster = 65, minmatch = 20.

DNA replication and repair

DLP1 and DLP2 have seven and eight proteins, respectively, identified to be involved in DNA replication and repair at the same gene position: DNA polymerase (gp27), replicative clamp (gp28), RecB exonuclease (gp31 - DLP2 only), RecA (gp34), MazG (gp36), replicative primase/helicase (gp39), Vsr endonuclease (gp48) and dCMP deaminase (gp54 in DLP1, gp55 in DLP2) (Table 2-3 and Table 2-4; Figure 2-5). Three and two additional proteins were assigned putative functions following bioinformatics analysis, in DLP1 and DLP2, respectively, based on their high percent identity to known proteins: DEAD box helicase (gp30), RecB exonuclease (gp31 - DLP1 only) and a transcriptional regulator (gp37) (Table 2-3 and Table 2-4; Figure 2-5).

BLASTp analysis for DLP1 and DLP2 gp27 shows it is 99 % identical to a putative DNA polymerase in PA25. The replicative clamp of PA25 shares a 99 % identity to gp28 of DLP1 and 95 % identity to gp28 of DLP2. In DLP1, gp31 is 100 % identical to ORF030 in PA73, though its putative function was assigned due to its 95 % identity to the RecB exonuclease of Ab26. Gp31 in DLP2 has 99 % identity to the RecB exonuclease of Ab26. HHpred analysis of the gp31 protein for DLP1 and DLP2 revealed the proteins are similar to the exonuclease of the λ Red recombination complex (DLP1: 99 % probability, E-value of 3.8e-17; DLP2: 99 % probability, E-value 5.6e-17)[192]. The RecA protein of Ab26 shares 99 % and 100 % identity to the gp34 proteins of DLP1 and DLP2, respectively. The protein gp37 of DLP1 (100 % identity to gp35 of PA73) and DLP2 (99 % identity to hypothetical protein in Ab26) have been assigned the putative function of transcriptional regulator due to their identity to the transcriptional regulator of the *Burkholderia* phage KL1[181]. The protein gp38 for both DLP1 and DLP2 shares 99 % identity to the replicative primase/ helicase found in PA25. Both DLP1 and DLP2 contain a Vsr endonuclease (gp48), although gp48 of DLP1 shares the highest identity to the KL1 Vsr endonuclease, whereas gp48 of DLP2 shares the highest identity to the Vsr endonuclease encoded by Ab26 (Table 2-3 and Table 2-4). The dCMP deaminase (gp54 of DLP1 and gp55 in DLP2) of both phages is most similar to the dCMP deaminase of PA25 (98 and 97 % identity, respectively). Protein gp30 of DLP1 and DLP2 is 100 % identical to gp29 of PA73, but a putative function has been assigned in both phages, as gp30 is 99 % identical to the DEAD box helicase protein of Ab26 for both phages. DEAD box helicases are vital to cellular functions as they fold RNA molecules into their correct secondary structures and realign RNA-protein interactions with the use of ATP[193].

A predicted protein of interest in DLP1 and DLP2 is MazG, which is encoded by gp36 in DLP1 (89 % identity to MazG of PA25), and gp36 in DLP2 (98 % identity to MazG of Ab26). During times of stress in bacteria, the unusual nucleotides pppGpp and ppGpp begin to accumulate, synthesized by the proteins SpoT and RelA respectively [194]. Amino acid starvation activates RelA to synthesize ppGpp, whereas other bacterial stressors such as carbon or nitrogen starvation triggers SpoT to synthesize pppGpp [194,195]. The pppGpp nucleotide can be converted into ppGpp though the enzyme GppA phosphatase[194]. Both of these unusual nucleotides are involved in the global response to stressful conditions within the bacteria, though ppGpp is a more potent regulatory nucleotide for growth inhibition [194,196]. MazG fits into this regulatory pathway by depleting the accumulated ppGpp, thus reducing growth inhibition[197]. The action of phage-encoded MazG has been of interest to researchers, as many marine phages have been found to encode MazG homologs[198]. It has been speculated that phage-encoded MazG operates to reduce the ppGpp pool within stationary-phase infected cells[199], thus enhancing propagation of phage progeny in bacterial cells growing in nutrient limiting conditions. The host bacterium for DLP1 and DLP2, *S. maltophilia*, has been isolated from nutrient-limited environments, such as ultrapure and deionized water [24,64]. The presence of MazG in DLP1 and DLP2 could potentially offer a competitive advantage over MazG-deficient phages replicating in stationary phase *S. maltophilia*.

Phage relatedness

The two *S. maltophilia* phages DLP1 and DLP2 differ from each other based upon RFLP analysis, DNA comparison analysis, protein to protein comparison analysis, the presence of insertions/deletions (indels), genetic synteny, as well as the phenotypic differences presented, which include different host ranges and the timing of plaque formation. Based upon these analyses, which include changes to both structural and replication genes and their predicted gene products, we conclude that they are significantly different enough in genetic content and biology to be considered individual phages and not merely variants of one another. There are sufficient genomic, proteomic and biological differences that, although they are related phages, DLP1 and DLP2 are not (or are no longer) close variants of each other. These differences include 1369 base pair changes and 157 gaps required to align the DNA, and three genomic locations where DLP1 and DLP2 have acquired completely different genes, which originate from entirely different

source (Table 2-5). In DLP1, ORF 32 encodes a protein of 175 amino acids which does not share any significant identity to proteins in the NCBI database, whereas in DLP2, ORF 32 encodes a 53 amino acid protein similar to a protein encoded by phage PA25. In DLP1, ORF 39 encodes a 66 amino acid protein which does not have significant results with BLASTp. This lack of results with BLASTp is also seen for the ORF 39 product of DLP2, which is a 150 amino acid long protein. In DLP2, ORF 55 encodes a 55 amino acid protein that has is similar to a protein encoded by both PA25 and PA73, whereas DLP1 has no coding DNA in this region of its genome. Besides these obvious differences, and even though these phages exhibit high average identity across their entire genomes and share almost all proteins, it is still possible that the differences within protein sequences may be associated with the functional differences displayed during bacterial host infection. For example, there are proteins that differ significantly between DLP1 and DLP2 (i.e. a large number of gaps are required to complete alignment), such as gene products 13, 21, 33, 48, 51 and 54 (encoding a hypothetical protein, a tail tape measure structural protein, a hypothetical protein, a Vsr endonuclease replication protein, a hypothetical protein and a dCMP deaminase replication protein, respectively) (Table 2-5).

Similar to bacterial relatedness, we suggest that phage relatedness is an arbitrary ideal, and that there are no set guidelines as to what constitutes a phage variant versus what is a different but related phage. In order of relatedness, it is clear that DLP1 is most related to DLP2, and then in order of decreasing relatedness, *Pseudomonas* phages PA25, PA73, Ab26 (Table 2-6) and finally *Burkholderia* phage KL1 (Table 2-5). Together, they are similar enough to be considered as a *Siphoviridae* sub-family, but how would one delineate them as variants of the same phage versus related phages of a continuum? For example, DLP2 is more related to PA25 circa ORF 32, and more related to PA25 and PA73 circa ORF 55, than DLP1, even though DLP2 shares the highest identity to DLP1. Therefore, how much genetic, proteomic and biological differences must be evident before phages are separated into different "species"? Based upon the biological differences (including host range and plaque formation differences), the significant genetic alterations (including the presence of "indels"), and the protein level differences (highly variable but related protein sequences) presented, we conclude that each of these phages are different but related phages. These analyses confirm the idea that, although the phage genome DNA sequences are syntenic, significant changes have occurred between every member of this

sub-family of phages, which is also reflected in the biological differences exhibited by phages DLP1 and DLP2.

Table 2-5: BLASTp comparison of DLP1 gene products to phage DLP2 and *Burkholderia* phage KL1

DLP1		**DLP2**					**KL1**				
Predicted protein		AA	ID	Sim	Exp	Gaps	AA	ID	Sim	Exp	Gaps
gp	AA										
1	84	84	81/ 84	81/ 84	3e-57	0	88	65/ 84	77/ 84	3e-40	0
2	88	88	84/ 88	85/ 88	4e-60	0	98	35/ 86	53/ 86	1e-16	0
3	181	181	178/181	179/181	1e-136	0					
4	101	101	100/101	101/101	2e-71	0	105	43/ 105	61/ 105	5e-14	8
5	93	93	91/ 93	91/ 93	6e-67	0	87	37/ 72	43/ 72	2e-12	9
6	162	162	161/162	161/162	4e-121	0	163	92/ 153	113/153	6e-55	2
7	531	531	524/531	528/531	0	0	546	433/513	467/513	0	3
8	501	501	484/499	490/499	0	0	498	417/486	443/486	0	0
9	364	364	362/364	364/364	0	0	371	268/365	314/365	0	1
10	239	239	235/239	237/239	3e-173	0	235	192/235	214/235	8e-133	0
11	325	325	322/325	324/325	0	0	324	291/325	308/325	0	1
12	134	134	133/134	133/134	3e-94	0	138	96/139	101/139	8e-49	6
13	133	123	93/ 133	106/133	5e-55	13	125	88/134	100/134	1e-45	13
14	172	172	155/172	164/172	1e-117	0	172	139/172	154/172	1e-98	0
15	126	126	120/126	122/126	1e-89	0	124	73/124	89/124	2e-43	2
16	151	151	145/151	147/151	6e-111	0	152	102/147	121/147	3e-71	0
17	511	511	507/511	508/511	0	0	517	403/514	442/514	0	5
18	142	142	140/142	141/142	7e-105	0	142	112/142	129/142	6e-79	0
19	118	114	112/114	114/114	4e-84	0	264	92/118	103/118	5e-52	0
20	144	144	144/144	144/144	3e-107	0	144	115/144	128/144	4e-77	0
21	1175	1187	1063/ 1214	1085/ 1214	0	66	1272	633/1117	764/1117	0	72
22	320	320	315/320	318/320	0	0					
23	321	321	306/321	316/321	0	0					
24	570	570	555/570	559/570	0	0	563	179/504	281/504	9e-98	6
25	274	274	272/274	274/274	0	0	273	136/276	178/276	7e-85	7
26	813	813	799/813	805/813	0	0	803	382/812	512/812	0	13
27	683	683	677/683	681/683	0	0	689	570/683	623/689	0	7
28	380	380	368/380	374/380	0	0	386	288/389	325/389	0	16
29	118	76	70/ 71	70/ 71	4e-50	0					
30	551	551	551/551	551/551	0	0	551	498/551	528/551	0	0
31	365	365	346/365	352/365	0	0	332	238/368	271/368	6e-167	0
32	175										
		53 (gp32)									

33	248	251	200/234	203/234	3e-103	24	261	111/138	123/138	4e-74	4
34	238	238	237/238	238/238	0	0	238	216/238	228/238	2e-159	0
35	146	146	143/146	145/146	2e-103	0	169	73/ 146	86/ 146	6e-32	9
36	185	185	160/185	169/185	1e-118	0	179	127/179	141/179	3e-75	1
37	68	68	68/ 68	68/ 68	1e-48	0	62	31/ 61	42/ 61	6e-13	0
38	773	773	763/773	769/773	0	0	769	656/773	709/773	0	4
39	66 NH										
40	49	49	38/ 49	43/ 49	5e-28	0					
41	59	63	58/ 59	59/ 59	5e-41	0					
42	98	98	96/ 98	96/ 98	8e-72	0	98	68/ 97	80/ 97	5e-43	0
43	59	59	52/ 59	57/ 59	1e-37	0					
44	80	80	78/ 80	78/ 80	4e-57	0	78	39/ 80	51/ 80	2e-14	2
45	73	73	73/ 73	73/ 73	5e-51	0					
46	124	124	121/124	121/124	5e-85	0	133	50/ 123	73/ 123	4e-10	3
47	73	73	72/ 73	72/ 73	3e-51	0	162	125/161	134/161	1e-83	1
48	163	178	135/181	140/181	3e-91	21	162	125/161	134/161	1e-83	1
49	71	71	70/ 70	70/ 70	9e-51	0	66	47/ 62	54/ 62	4e-27	0
50	59	59	57/ 59	57/ 59	4e-41	0	62	25/ 56	35/ 56	.006	5
51	99	101	82/ 101	88/ 101	8e-55	2					
52	96	96	94/ 96	96/ 96	5e-67	0					
53	77	77	76/ 77	77/ 77	8e-58	0	78	50/ 76	62/ 76	2e-32	0
54	155	145	139/145	141/145	5e-106	0	161	112/155	121/155	4e-73	3
55	127	127 (gp56)	127/127	127/127	5e-95	0	124	54/ 127	76/ 127	8e-19	10
56	69	69 (gp57)	66/ 69	66/ 69	8e-49	0	68	29/ 69	45/ 69	1e-14	1
57	190	190 (gp58)	187/190	189/190	1e-137	0	184	124/161	185/161	2e-61	2

Abbreviations: gp – gene product; AA – amino acid; ID – AA identity; Sim – AA similarity; Expect – expected number of chance matches in a random model; Gaps – number of gaps added to increase alignment; NH – no homolog in database. Empty cells indicate no homolog in phages DLP2 or KL1.

Bioinformatic analysis of central tail hub

Investigation into the receptor used by both DLP1 and DLP2 was conducted by the graduate student, Jaclyn McCutcheon, by screening a mini-Tn5-*luxCDABE P. aeruginosa* PA01 mutant library[200] to identify any mutants resistant to DLP1 infection. Mutants in type IV pilus (T4P) components and regulatory factors led to the discovery of the DLP1 receptor. Further work in *S. maltophilia* D1585 confirmed PilA and PilE mutants were resistant to both DLP1 and DLP2 infection, indicating both phages use the T4P to enter their host cell. Investigation of the central tail hub proteins of DLP1 and DLP2 revealed they both contain the Phage-tail_3 domain (Pfam 13550) which has been implicated in host specificity[201]. A literature search of known type

IV pili binding *Siphoviridae* phages and subsequent CD-domain search of each genome led to the discovery that all of the known type IV pili-binding phages encode the Phage-tail_3 domain protein and do not encode any proteins with tail fiber domains. The TEM images of DLP1 and DLP2 (Figure 2-2) also suggest a lack of tail fibers. Siphoviruses lacking tail fibers is not novel[202-204]. For example, the temperate phage J-1 was isolated from an abnormal fermentation of *Lactobaccilus casei* and noted to lack tail fibers[205-207]. Lack of tail fibers might benefit the phage if the receptor site is excessively small, allowing for a more streamlined profile. If the phage uses pili-retraction to be brought to their receptor, they could potentially use their baseplate to attach to the pilus for eventual retraction to secondary receptor.

A MUSCLE alignment of the central tail hub proteins encoded by the known type-IV pili binding phages shows high percent identity between the majority of *P. aeruginosa* phages, though B3 and JBD23 appear more divergent (Table 2-7). DLP1 and DLP2 show high homology to each other, but are divergent from the known T4P phages, sharing a maximum of 25.5 % identity to the *S. maltophilia* phage DLP4 for both DLP1 and DLP2. It has been experimentally confirmed in the lab by the graduate student Jaclyn McCutcheon that DLP4 uses the type IV pili for infection, though this data is not yet published. A comparison of the DLP1 and DLP2 central tail hub proteins to the highest ranking BLASTp hits was conducted against the virus database and a MUSCLE alignment was performed using all of the proteins. The resulting table shows high homology between DLP1, DLP2 and the *P. aeruginosa* phages 73, vB_PaeS_C1, vB_Pae-Kakheti25 (PK25), vB_Pae_PS9N, and vB_PaeS_SCH_Ab26 (AB26) (Table 2-8). The isolation and characterization of *Burkholderia* phage KL1 was documented in 2012, but the receptor remains to identified[181]. The 46.5 % identity shared between DLP1, DLP2 and KL1 suggests the receptor for KL1 in *Burkholderia* may be the type IV pili, though this requires further investigation.

Table 2-6: BLASTp alignment of DLP1 against *P. aeruginosa* phages PA25, PA73, and Ab26.

DLP1		**PA25**					**PA73**					**Ab26**				
Predicted protein																
gp[1]	AA [2]	AA	ID[3]	Sim[4]	Exp[5]	Gaps	AA	ID	Sim	Exp	Gaps	AA	ID	Sim	Exp	Gaps
1	84	84	83/ 84	83/ 84	4e-51	0	84	84/ 84	84/ 84	1e-51	0					
2	88	88	87/ 88	87/ 88	5e-55	0	88	85/ 88	86/ 88	4e-53	0	88	81/ 88	84/ 88	4e-51	0
3	181	181	178/ 181	179/ 181	8e-129	0	181	178/ 181	179/ 181	2e-128	0	181	177/ 181	179/ 181	2e-127	0
4	101	101	100/101	101/ 101	5e-63	0	101	99/ 101	101/101	1e-62	0	101	100/101	101/101	4e-63	0
5	93	93	93/ 93	93/ 93	3e-60	0						93	91/ 93	91/ 93	1e-58	0
6	162	162	161/162	161/ 162	6e-113	0	162	162/162	162/162	4e-114	0	162	161/162	161/162	6e-113	0
7	531	531	529/531	530/531	0	0	531	529/531	530/531	0	0	531	525/531	528/531	0	0
8	501	501	495/501	496/501	0	0	501	488/501	493/501	0	0	510	488/500	493/500	0	0
9	364	364	359/364	364/364	0	0	364	364/364	364/364	0	0	364	363/364	364/364	0	0
10	239	239	239/239	239/239	2e-168	0	239	238/239	238/239	1e-167	0	239	235/239	239/239	2e-166	0
11	325	325	323/325	324/325	0	0	325	323/325	324/325	0	0	324	303/325	310/325	0	1
12	134	134	134/134	134/134	1e-86	0	134	134/134	134/134	1e-86	0	130	128/134	129/134	4e-81	4
13	133	133	92/ 131	103/131	4e-45	13	123	92/ 131	103/131	2e-45	13	123	94/133	104/133	4e-47	13
14	172	172	155/172	164/172	2e-109	0	172	154/172	163/172	8e-109	0	172	159/172	164/172	1e-111	0
15	126	126	124/126	124/126	2e-84	0	126	123/126	123/126	1e-83	0	126	120/126	123/126	8e-82	0
16	151	151	146/151	147/151	5e-103	0	151	146/151	147/151	5e-103	0	151	146/151	147/151	6e-103	0

17	511	511	508/511	509/511	0	0	511	508/511	510/511	0	0	511	491/511	497/511	0	0
18	142	142	142/142	142/142	1e-97	0	142	142/142	142/142	1e-97	0	142	140/142	141/142	8e-97	0
19	118	264	117/118	118/118	1e-78	0	118	118/118	118/118	8e-80	0	118	117/118	118/118	2e-79	0
20	144	144	144/144	144/144	5e-99	0	144	144/144	144/144	5e-99	0	144	144/144	144/144	5e-99	0
21	1175	1204	1162/1204	1170/1204	0	29	1204	1161/1204	1171/1204	0	29	1204	1146/1204	1167/1204	0	29
22	320	320	279/320	296/320	0	0	320	259/320	282/320	0	2	320	257/320	279/320	0	2
23	321	317	211/326	244/326	3e-132	14	303	178/322	221/322	3e-106	20	375	173/322	218/322	3e-102	20
24	570	571	548/571	558/571	0	1	568	500/570	527/570	0	2	568	500/570	525/570	0	2
25	274	274	271/274	274/274	0	0	274	270/274	273/274	0	0	274	269/274	273/274	0	0
26	813	813	801/813	805/813	0	0	813	809/813	810/813	0	0	813	796/813	805/813	0	0
27	683	683	680/683	681/683	0	0	683	680/683	682/683	0	0	683	677/683	681/683	0	0
28	380	380	379/380	379/380	0	0	380	378/380	379/380	0	0	387	375/380	378/380	0	0
29	118	75	70/ 71	70/ 71	1e-41	0	118	115/118	116/118	7e-78	0					
30	551	551	550/551	550/551	0	0	551	551/551	551/551	0	0	554	550/551	551/551	0	0
31	365	365	365/365	365/365	0	0	365	365/365	365/365	0	0	365	347/365	352/365	0	0
32	175															
33	248	248	246/248	248/248	2e-165	0	248	245/248	247/248	1e-164	0	235	131/132	132/132	5e-88	0
34	238	238	237/238	238/238	2e-173	0	238	237/238	238/238	2e-173	0	238	237/238	238/238	3e-173	0
35	146	146	145/146	145/146	2e-96	0	146	145/146	145/146	2e-96	0	238	142/146	144/146	1e-94	0

36	185	185	165/185	173/185	7e-115	0	185	164/185	172/185	2e-113	0	185	159/185	169/185	7e-115	0
37	68	62	62/ 62	62/ 62	2e-35	0	62	62/ 62	62/ 62	2e-35	0	68	67/ 68	68/ 68	2e-39	0
38	773	773	770/773	771/773	0	0	773	769/773	770/773	0	0	780	763/773	769/773	0	0
39	66 NH															
40	49															
41	59						63	55/ 59	57/ 59	5e-29	0					
42	98	91	25/ 60	38/ 60	6e-08	3	98	96/ 98	96/ 98	2e-63	0	91	25/ 60	38/ 60	6e-08	3
43	59	55	51/ 55	54/ 55	3e-28	0						67	53/ 59	55/ 59	2e-29	0
44	80	80	76/ 80	77/ 80	4e-47	0	80	77/ 80	78/ 80	3e-48	0	80	77/ 80	78/ 80	3e-48	0
45	73															
46	124	124	123/124	124/124	7e-79	0	124	120/124	121/124	1e-69	0	124	117/124	120/124	1e-66	0
47	73	73	71/ 73	72/ 73	1e-42	0						74	53/ 73	60/ 73	1e-28	1
48	163	178	135/181	140/181	4e-83	21	179	135/181	140/181	4e-83	21	181	139/181	146/181	6e-92	18
49	71	67	67/ 67	67/ 67	8e-40	0						131	70/ 71	71/ 71	1e-42	0
50	59	59	56/ 59	56/ 59	2e-31	0										
51	99	101	82/ 101	87/ 101	6e-46	2	101	82/ 101	88/ 101	2e-46	2					
52	96	96	95/ 96	96/ 96	3e-59	0	96	95/ 96	96/ 96	3e-59	0	126	93/ 96	94/ 96	4e-58	0
53	77	77	77/ 77	77/ 77	2e-49	0	77	76/ 77	77/ 77	8e-49	0	80	75/ 77	77/ 77	2e-48	0
54	155	155	152/155	153/155	4e-107	0	155	149/155	151/155	1e-105	0	155	150/155	151/155	3e-105	0
55	127	127	125/127	127/127	7e-86	0	127	127/127	127/127	9e-87	0	129	89/ 89	89/ 89	2e-57	0
56	69	69	65/ 69	66/ 69	7e-40	0	94	66/ 69	66/ 69	7e-41	0	94	65/ 69	66/ 69	2e-40	0
57	190	190	189/190	190/190	3e-131	0	190	189/190	190/190	3e-131	0	190	182/190	185/190	2e-125	0

Abbreviations: gp – gene product; AA – amino acid; ID – AA identity; Sim – AA similarity; Expect – expected number of chance matches in a random model; Gaps – number of gaps added to increase alignment; NH – no homolog in database. Empty cells indicate no homolog in phages DLP2 or KL1.

Table 2-7: MUSCLE alignment of the central tail hub proteins (Phage-tail_3 domain; Pfam 13550) of known type-IV pili binding *Siphoviridae* phages from literature. Values indicate percent identity between each protein.

	DLP1	DLP2	DLP4	B3	D3112	JBD5	JBD26	JBD30	JBD69	JBD93	MP22	MP29	MP42	Salvo[a]	Sano[a]	JBD23	phiCbK[b]
DLP1		98.3	25.5	22.6	23.4	23.4	23.4	23.9	23.5	23.9	23.7	23.8	23.3	24.2	24.9	8.6	11.4
DLP2	98.3		25.5	23.4	23.8	23.8	23.8	24.3	23.9	24.3	24.1	24.3	23.7	24.2	25	8.6	11.2
DLP4	25.5	25.5		30.3	27.9	28.3	27.8	27.8	27.9	27.6	27.2	28.1	27.6	26.7	27.9	9.9	12.5
B3	22.6	23.4	30.3		50.9	50.6	50.5	51	50.7	50.7	50.6	50.9	50.1	22.5	23.1	9.3	12.4
D3112	23.4	23.8	27.9	50.9		97	97.1	98.1	98.4	98.1	97.4	96.5	96.1	21.4	21.3	11.6	14.4
JBD5	23.4	23.8	28.3	50.6	97		97.4	96.5	96.5	96.1	97.3	98.9	95.1	21.8	21.4	11.6	14
JBD26	23.4	23.8	27.8	50.5	97.1	97.4		96.5	96.1	96.3	95.9	97.1	94.3	21.3	20.8	11.4	13.9
JBD30	23.9	24.3	27.8	51	98.1	96.5	96.5		97	98.9	98.4	96.5	96.2	21.9	21.6	11.8	14.6
JBD69	23.5	23.9	27.9	50.7	98.4	96.5	96.1	97		97	97.1	96.2	95.2	21.8	21.2	11.3	14.4
JBD93	23.9	24.3	27.6	50.7	98.1	96.1	96.3	98.9	97		98	96.3	96.3	21.9	21.6	12	14.4
MP22	23.7	24.1	27.2	50.6	97.4	97.3	95.9	98.4	97.1	98		97.3	96.6	22	21.6	12	14.4
MP29	23.8	24.3	28.1	50.9	96.5	98.9	97.1	96.5	96.2	96.3	97.3		94.8	21.8	21.3	11.7	14
MP42	23.3	23.7	27.6	50.1	96.1	95.1	94.3	96.2	95.2	96.3	96.6	94.8		21.7	21.3	11.6	14.1
Salvo	24.2	24.2	26.7	22.5	21.4	21.8	21.3	21.9	21.8	21.9	22	21.8	21.7		73.7	7.5	11.6
Sano	24.9	25	27.9	23.1	21.3	21.4	20.8	21.6	21.2	21.6	21.6	21.3	21.3	73.7		7.6	12.3
JBD23	8.6	8.6	9.9	9.3	11.6	11.6	11.4	11.8	11.3	12	12	11.7	11.6	7.5	7.6		6.7
phiCbK	11.4	11.2	12.5	12.4	14.4	14	13.9	14.6	14.4	14.4	14.4	14	14.1	11.6	12.3	6.7	

B3, D3112, JBD*, and MP* are *Pseudomonas* phages. [a] Salvo and Sano are *Xylella* phages. [b] PhiCbK is a *Caulobacter* phage.

Table 2-8: MUSCLE alignment of DLP1 and DLP2 phage central tail hub proteins against the top ten BLASTp results in the virus database. All proteins have a Phage-tail_3 domain (Pfam13550).

Bacteriophage	Accession	% identity to Pfam13550	
		DLP1	DLP2
Stenotrophomonas phage DLP1[a]	AKI28788.1	-	98.3
Pseudomonas phage 73	YP_001293432	99.5	98.5
Pseudomonas phagevB_PaeS_C1	AVJ48095	98.8	99.3
Pseudomonas phage vB_Pae-Kakheti25	YP_006299890	98.5	99.1
Pseudomonas phage vB_Pae_PS9N	AIW01689	98.4	98.4
Stenotrophomonas phage DLP2[a]	AKI28730.1	98.3	-
Pseudomonas phage vB_PaeS_SCH_Ab26	YP_009044360	97.9	97.4
Pseudomonas phage PaMx42	YP_009205621	69.3	69.7
Burkholderia phage KL1	YP_006560777	46.8	46.8
Xylella phage Sano[a,b]	AHB12068	29.5	29.4
Xylella phage Salvo[a,b]	AHB12243	29	28.8

[a] Experimentally confirmed as pili-binding phages. [b] One tail fiber gene annotated, but no tail fiber homologues using nucleotide sequence for CD-Search against database CDD v3.16-50369 PSSMs with the expected E-value threshold of 0.01, and composition-based statistics adjustment checked.

CONCLUSIONS

Although relatively rare, the incidence of phage broad host range specificity at the genera level is being increasingly studied in the food production industry, mainly with the *Salmonella* and *Escherichia* genera[208-211]. This study is the first to identify and characterize phages capable of infecting pathogenic bacteria across taxonomic orders. DLP1 and DLP2 are closely related phages that share high identity to *P. aeruginosa* phages vB_Pae-Kakheti25, vB_PaeS_SCH_Ab26, and PA73 and lesser identity to *Burkholderia* phage KL1. Phage DLP1, possessing a 42,887 bp genome, is predicted to encode 57 proteins and exhibits a delayed plaque development phenotype. Unlike DLP1, phage DLP2 exhibits normal plaque development, but possesses a similar genome 42,593 bp in length. The cause of the delayed plaque development in phage DLP1 is unknown, but genomic comparison suggests that gene variants encoded by or acquired by DLP1 may contribute to the observed lysis phenotype differences. Another possibility is a protein encoded by both phages involved in lysis timing is mutated in DLP1 resulting in the delay. The use of phage therapy may be one of the best treatment options for

otherwise untreatable drug resistant bacterial infections [212-214]. The genomic characterization of broad-host range phages such as DLP1 and DLP2 is the first step towards developing an effective phage therapy strategy for *S. maltophilia*.

ACKNOWLEDGEMENTS

The authors would like to thank Arlene Oatway from the University of Alberta Department of Biological Sciences Advanced Microscopy Facility for assistance with electron microscopy, and members of the Dennis lab for helpful scientific discussions. The authors thank the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR, Vancouver, BC) and The Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services, for gifts of bacterial strains. JJD gratefully acknowledges operating grant funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Chapter 3

Temperate *Stenotrophomonas maltophilia* bacteriophage DLP3 lysogeny causes lysogenic conversion of host strain D1571

OBJECTIVES

The objectives of this research were to isolate and characterize the bacteriophage DLP3 though analysis of its lifestyle, morphology, and genomic composition.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Initial phage isolation was accomplished with five clinical *Stenotrophomonas maltophilia* strains (D1585, D1571, D1614, D1576, and D1568) from the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR; Vancouver, B.C.). An additional 22 clinical isolates were obtained from the Provincial Laboratory for Public Health – North (Microbiology), Alberta Health Services for host range analysis. Strains were grown aerobically at 30 ˚C on Luria-Bertani solid media until single colonies were visible (16 – 36 h) or in LB broth with shaking at 225 RPM.

Bacteriophage isolation, propagation, and host range

The phage DLP3 (vB_SmaS_DLP_3) was isolated from soil collected in Edmonton, Alberta, Canada using the *S. maltophilia* strain D1571. No plants were associated with the soil sample. Approximately 10 ml of soil was mixed with 10 ml of LB broth, 1 ml of modified suspension media (SM), and 100 µl of a D1571 overnight culture. The slurry was incubated overnight at 30 ˚C with shaking. The supernatant was filter-sterilized with a Millex-HA 0.45 µm syringe-driven filter (Millipore, Billerica, MA) and stored at 4 ˚C[163]. A single plaque was picked to propagate a working-stock solution for analysis using top-agar overlays. Briefly, 100 µl of overnight D1571 culture and 100 µl DLP3 stock (~10^9 PFU/ml) were mixed and incubated for 5 min at room temperature then added to 3 ml of 0.7 % LB top agar. The mixture was poured onto an LB plate and incubated for 18 h at 30 ˚C. The top agar of plates showing confluent lysis was scraped into a 50 ml Falcon tube. A 3 ml aliquot of SM was added for each plate scraped, and the slurry was shaken for 1 min followed by centrifugation (5 min at 10,000 x g) and filter-sterilization. Host range analysis was performed using serially-diluted DLP3 lysate into SM. A

10 µl aliquot of each concentration was spotted in triplicate onto a plate containing one of 27 *S. maltophilia* strains in a top-agar overlay and incubated overnight at 30 ˚C.

Electron microscopy

Phage lysate for electron microscopy was prepared using LB plates, and top agar made with agarose and filter sterilized using a 0.22 µm filter. A carbon-coated copper grid was overlaid with 10 µl of phage lysate for 2 min then stained with 4 % uranyl acetate for 30 s. A Philips/FEI (Morgagni) transmission electron microscope (TEM) with charge-coupled device camera at 80 kV (University of Alberta Department of Biological Sciences Advanced Microscopy Facility) was used to obtain TEM images. The capsid diameter, tail length and tail width of ten virions were measured using ImageJ and averages calculated using Microsoft Excel.

Phage DNA isolation, sequencing, and RFLP analysis

Genomic DNA was isolated from a high-titer DLP3 stock (10^9 PFU/ml). Lysate was clarified by spinning 10,000 x g for 10 min, and the supernatant was treated with 100 µl 100x DNase I buffer (1 M Tris–HCl, 0.25 M $MgCl_2$, 10 mM $CaCl_2$), 10 µl DNase I (Thermo Scientific, Waltham, MA), and 6 µl RNase (Thermo Scientific) and incubated 1 h at 37 ℃. A 400 µl aliquot of 0.5 M EDTA (pH 8.0), SDS (final concentration of 2 %) and Proteinase K (final concentration of 400 µg/ml) was added followed by incubation at 55 ˚C overnight. A ½ volume of 6 M NaCl was added, and the solution was vortexed at high speed for 30 s followed by centrifugation at 17,900 x g for 30 min. The supernatant was transferred to a fresh tube with an equal volume of 100 % isopropanol and stored at -20 ˚C for at least 1 hour to overnight. The DNA was pelleted with centrifugation at 17,900 x g for 20 min at 4 ˚C followed by three 70 % ethanol washes. The pellet was dried at room temperature and resuspended in nuclease-free water. Purity and concentrations of eluted DNA were checked with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA). DLP3 genomic (gDNA) DNA was sequenced using both Illumina and Pacific Biosciences technology. A Nextera XT library was generated for paired-end sequencing on MiSeq (Illumina) platform using MiSeq v2 reagent kit[215]. The DLP3 gDNA was also sequenced on a PacBio RS II platform at the Genome Quebec Innovation Center to determine what sequences of the DLP3 genome have DNA modifications.

Restriction fragment length polymorphism (RFLP) analysis was used with 15 FastDigest (Thermo Scientific) restriction enzymes: EcoRI, XbaI, BamHI, HindIII, KpnI, SmaI, SphI, PstI,

SacI, SaII, ApaI, ClaI, NdeI, SpeI, and XhoI. Restriction reactions were set up using 1 µl FastDigest enzyme, 2 µl FastDigest restriction buffer, 1 µg of phage DNA and nuclease-free water to bring the final volume to 20 µl. Reactions were separated on a 0.8 % (wt/vol) agarose gel in 1x TAE (pH 8.0).

Protein isolation and mass spectrometry

Isolation of DLP3 protein for SDS-PAGE analysis was accomplished following a protocol for the formation of ghost particles[216]. Sterile DLP3 lysate (~1 x 10^9) was clarified twice with 10,000 x g centrifugations and treated with nucleases following the DNA isolation protocol described above. After the incubation, an equal volume of 10 M LiCl was added, and the solution was incubated at 46 ˚C for 10 min, followed by 10-fold dilutions into sterile Milli-Q water. The released DLP3 gDNA was digested with an addition of 10 mM $MgCl_2$ and 50 U of RNase-free DNaseI per 1x10^{12} PFU. This solution was incubated overnight at 37 ˚C, followed by ultracentrifugation at 28,700 x g for 1.2 h. The supernatant was discarded, and pellets were resuspended with 100 µl SM. An aliquot of the sample was diluted in half with 2x Laemmli sample buffer (10 % [v/v] beta-mercaptoethanol [BME], 6 % [w/v] SDS, 20 % [v/v] glycerol, and 0.2 mg/ml bromophenol blue) and incubated 10 min at 99 ˚C.

An SDS-PAGE gel with a 4 % stack and a 7.5 % resolving portion was made with 40 % 37.5:1 acrylamide/bis-acrylamide solution (Bio-Rad) and fresh 10 % ammonium persulfate. The gel was loaded into a Mini-PROTEAN electrophoresis chamber (Bio-Rad) using 1x running buffer. A 6 µl aliquot of PageRuler Plus Prestained Protein Ladder (Thermo Scientific) was used as a molecular weight standard and 6 – 12 µl of DLP3 ghost particles in 1x sample buffer was loaded into the remaining wells. The gel was run at 180 kV for 75 min and placed in Coomassie R-250 stain for 1 h with gentle rocking. The gel was destained over 2 h, with the destaining solution replaced every 30 min. The gel was placed in a 50 ml Falcon tube with Milli-Q to transport the gel for mass spectrometry analysis at the Alberta Proteomics and Mass Spectrometry (APM) facility located at the University of Alberta.

In-gel trypsin digestion was performed on the samples. The lane was cut into 7 equal gel sections, destained twice in 100 mM ammonium bicarbonate/ acetonitrile (ACN) (50:50), reduced (10 mM BME–100 mM bicarbonate), and alkylated (55 mM iodoacetamide–100 mM bicarbonate). After dehydration, trypsin digestion (6 ng/µl) was allowed to proceed overnight at

room temperature. Tryptic peptides were extracted from the gel using 97% water–2% acetonitrile–1% formic acid followed by a second extraction using 50% of the initial extraction buffer and 50% acetonitrile. Fractions containing tryptic peptides were resolved and ionized using nanoflow high-performance liquid chromatography (HPLC) (Easy-nLC 1000; Thermo Scientific) coupled to a Q Exactive Orbitrap mass spectrometer (MS) (Thermo Scientific). Nanoflow chromatography and electrospray ionization were accomplished by using a Pico- Frit fused silica capillary column (ProteoPepII; C_{18}) with a 100-µm inner diameter (New Objective) (300 Å, 5 µm pore size). Peptide mixtures were injected onto the column at a flow rate of 3,000 nl/min and resolved at 500 nl/min using 75-min linear gradients of 4% to 40% (vol/vol) aqueous ACN with 0.2% (vol/vol) formic acid. The mass spectrometer was operated in data-dependent acquisition mode, recording high-accuracy and high-resolution Orbitrap survey spectra using external mass calibration, with a resolution of 35,000 and *m/z* range of 400 to 2,000. The 15 most intensely multiply charged ions were sequentially fragmented by HCD fragmentation. After two fragmentations, all precursors selected for dissociation were dynamically excluded for 60 s. Data were processed using Proteome Discoverer 1.4 (Thermo Scientific). The UniProt *Stenotrophomonas* database and all DLP3 proteins were searched using SEQUEST (Thermo Scientific). Search parameters included a precursor mass tolerance of 10 ppm and a fragment mass tolerance of 0.8 Da. Peptides were searched with carbamidomethyl cysteine as a static modification and oxidized methionine and deamidated glutamine and asparagine as dynamic modifications.

Determination of DLP3 lifestyle

Top agar overlay plates showing confluent lysis of D1571 by DLP3 were used to obtain resistant colonies. Briefly, 3 ml of SM was added to the plates, and a sterile glass rod was used to gently skim the agar. The SM was collected and placed into microcentrifuge tubes (MCT), then centrifuged at 5,000 x g for 5 min. The supernatant was discarded, and 1 ml of fresh SM was added to resuspend the pellet, followed by centrifugation at 5,000 x g for 5 min. This wash step was repeated three times in total. Following the final wash centrifugation, the supernatant was removed, and the pellet was resuspended in 500 µl LB broth. Cells were serially diluted with LB and plated on LB plates, then incubated at 30 ˚C for 16 h. Single colony isolates were selected for further study and tested for superinfection resistance using overnight cultures of every isolate

in a top agar overlay assay with DLP3. After an 18 h incubation at 30 ˚C, the plates were observed for plaque development. Single colony isolates without plaque development were retained for analysis.

Growth analysis of wild type D1571 and the DLP3 lysogen

Single colony triplicate overnight cultures of wild type D1571 and the lysogen D1571::DLP3 were grown in LB broth at 30 ˚C with shaking. Subcultures (1:100) for each sample were performed using LB broth, and each subculture was grown to an OD600 of ~0.32 at 30 ˚C with 225 RPM shaking. Subcultures were distributed in triplicate aliquots of 200 µl 96 well plates, and an LB broth control was included for each plate. The OD600 was then obtained for each plate using a Wallac 1420 VICTOR2 multilabel counter (PerkinElmer, Waltham, MA) at the following time points: 0, 2, 4, 6, and 8 h. The OD600 data were used to determine the growth rate (µ) with the established formula: $\log10\ N - \log10\ N_0 = (\mu/2.303)\ (t - t_0)$, whereby N_0 is the time zero (t_0) OD600 reading, and N is the final OD600 reading obtained at a specific time (t) in the experiment. Resulting data were analyzed with GraphPad Prism 7 (GraphPad Software Inc., San Diego, CA) to graph the growth curve and growth rate. Statistical analysis of the growth rate was performed in GraphPad Prism 7 using a two-way ANOVA with multiple comparisons.

Bioinformatic analysis of the DLP3 genome

A 96,852 bp contig assembled from the Illumina reads with SPAdes 3.8.0 was identified for further analysis. No gaps or ambiguous sites were found in the assembly, which has a mean coverage of 114 reads and Q40 of 93.8%. Prediction of open reading frames (ORFs) was accomplished with the GLIMMER plugin [172] for Geneious [217] using the Bacteria and Archaea setting, as well as GeneMarkS for phage [174]. Conserved domain searches were performed using CD-Search[175] with the CDD v3.16 – 50369 PSSMs database. Phyre[218], HHblits[219,220], and I-TASSER[221] were used to gain insights into possible functions of hypothetical proteins or to provide more support for putative functions. BLASTn and BLASTp were used to gain information on relatives based on genomic data and individual proteins respectively [176]. The NCBI non-redundant protein sequence and nucleotide collection databases (update dates for both: 2018/08/26) were used for the BLASTp and BLASTn searches respectively. BLASTp

results above 1.00E-03 were annotated as hypothetical proteins. tRNAs were identified using the general tRNA model with tRNAscan-SE software[222].

RESULTS AND DISCUSSIONS

Isolation, morphology, host range, and RFLP analysis

Bacteriophage DLP3 (vB_SmaS-DLP_3) was isolated from a patch of barren soil collected in Edmonton, Alberta, Canada using clinical *Stenotrophomonas maltophilia* strain D1571. Transmission electron microscopy (TEM) enabled the classification of DLP3 as a *Siphoviridae* of the B1 morphotype[223] due to the long, noncontractile tail averaging 202.2 ± 5.7 nm and isometric capsid with a length and width of 92.8 ± 4.1 and 84.0 ± 2.8 nm respectively (Figure 3-1). No tail fibers were observed in the TEM images. The host range of DLP3 against all 27 clinical *S. maltophilia* isolates reveals a broad tropism though the successful infection of 21 strains (Table 1-2). The restriction fragment length polymorphism analysis revealed DLP3 genomic DNA is resistant to the 15 restriction enzymes screened: EcoRI, XbaI, BamHI, HindIII, KpnI, SmaI, SphI, PstI, SacI, SaII, ApaI, ClaI, NdeI, SpeI, and Xhol. These results suggest DLP3 contains modified DNA, though the types of modifications are unknown at this time.

Genomic characterization

The DLP3 genome is 96,852 bp long with a 58.3 % global GC content. No low coverage or ambiguous regions were identified with the assembled contig, which has a mean coverage of 114 and a Q40 of 99.6 %. NCBI non-redundant protein sequences database searches (update date: 2018/08/26) shows DLP3 shares a high identity to the *Siphoviridae* phage vB_SmaS_DLP_5 (DLP5). DLP5 is the type species of the new genus *Delepquintavirus* and based on the genomic similarities between DLP3 and DLP5, DLP3 is also a member of this new genus. A detailed overview of the similarities between DLP3 and DLP5 are detailed in Chapter 5. Open reading frame calling with Glimmer and GeneMarkS identified a total of 148 coding domain sequences (CDS) covering 95 % of the genome (Figure 3-2 and Table 3-1). DLP3 encodes five tRNA genes with different specificities: Tyr (GTA), Sup (CTA), Ser (GCT), Ile (GAT), and Glu (TTC). A total of 97 proteins could not be assigned functions due to lack of significant results from both BLASTp and CD-Search.

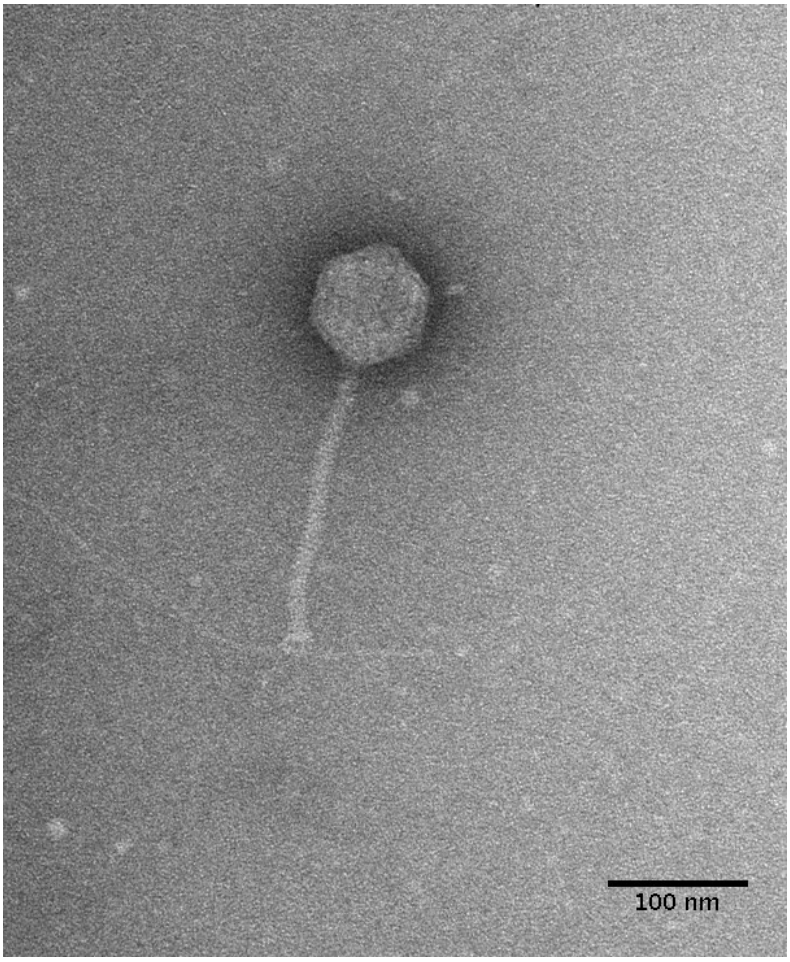

Figure 3-1: DLP3 *Siphoviridae* morphology. Phage lysate was applied to a carbon-coated copper grid and stained with 4 % uranyl acetate. Transmission electron micrographs were obtained at 180,000 x magnification. A *S. maltophilia* D1571 pili is shown to be interacting with the baseplate portion of the DLP3 tail. The averaged measurements for tail length, capsid length and width from ten virions is 202, 92, and 84 nm respectively.

The CD-Search did yield 37 DLP3 proteins with conserved domains predicted (Table 3-2). Three of the domains identified are domains of unknown function (DUF) which tended to be distributed throughout Gram-negative bacteria, and to a lesser extent Gram-positive, according to the species distribution for each DUF in pfam[224]. There are six conserved domains (CD) identified which are involved in virion morphogenesis: phage portal protein superfamily; gp1, phage capsid family; gp6, phage tail proteins; gp10 and 24, laminin G; gp28, and tape measure protein domain; gp17. Nine proteins with domains involved in DNA replication and repair identified with the CD-Search include two ParB domains; gp2 and 104, two helicases; gp49 and 54, Holliday junction resolvase; gp60, RecA recombinase; gp57, topoisomerase

primase; gp51, DNA ligase; gp130, and DNA polymerase A; gp145. The remaining CD results appear to be quite diverse in their functions such as the SpoVK family domain involved in sporulation (gp38)[225], protein-tyrosine phosphatases (gp139 and 141) typically involved in signal transduction[226], and a membrane-associated serine protease of the rhomboid family (gp93). One CD identified that is of particular interest is the glycosyltransferase domain of gp102. T-even bacteriophages have been shown to use glycosyltransferases for DNA modification by linking a glycosyl group to hydroxymethyl-cytosine, thus protecting the DNA against digestion by bacterial restriction systems[227]. Another role for glycosyltransferases within bacteriophages is highlighted by some *Shigella* phages which have been shown to seroconvert their host by modifying the O-antigen polysaccharides to prevent infection of the bacteria by other O-antigen receptor phages[228]. The specific glycosyltransferase family is RfaB, a protein involved in the assembly of the lipopolysaccharide (LPS) core of *Escherichia coli* K-12[229]. This result suggests DLP3 may use the glycosyltransferase to modify the host LPS, similar to the seroconverting *Shigella* phages, though this has yet to be confirmed experimentally.

Analysis of DLP3 structural proteins

To further investigate DLP3 morphology, phage structural proteins were analyzed by HPLC/MS and screened against DLP3 proteins and the *Stenotrophomonas* database in UniProt. The SEQUEST results from searching all DLP3 proteins identified 21 (Table 3-3), though only 11 proteins were classified as virion morphogenesis using BLASTp and CD-Search data (Figure 3-2). The most abundant protein isolated was the major capsid protein (gp6), which is the main structural component of a bacteriophage capsid (Figure 3-3)[230]. The second most abundant protein found was the portal protein (gp1), which forms the entry site for phage DNA to be packaged into the capsid by the large terminase. The portal protein also functions like a DNA-sensor, measuring the amount of DNA packaged into the capsid and signaling the large terminase to end genome-packaging once full[231](Figure 3-3). Higher concentrations of the gene products 14 and 18 were identified. Both gene products are hypothetical proteins without conserved domains and they do not have significant results with HHpred; therefore, their structural function is unknown (Table 3-2).

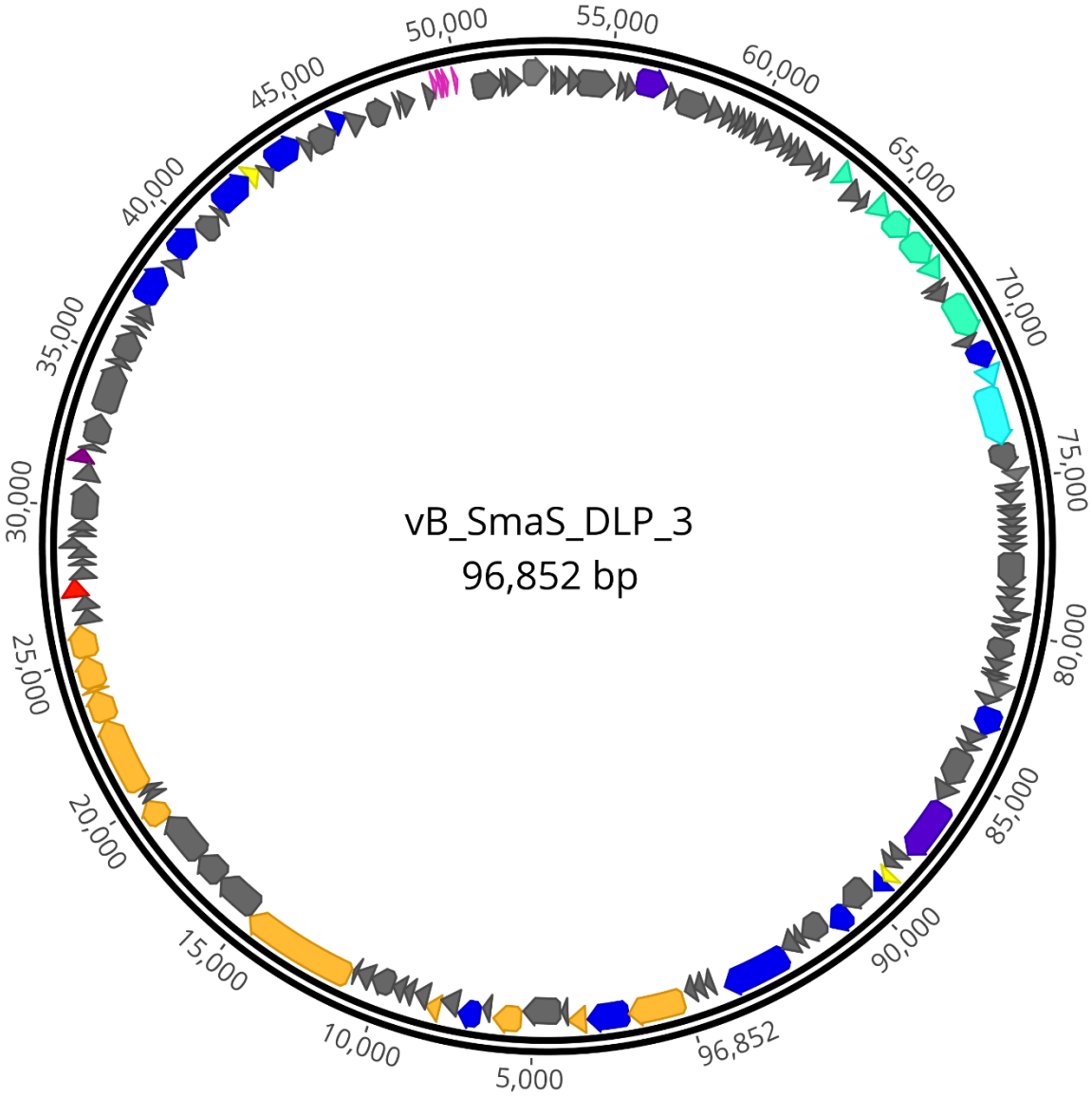

Figure 3-2: Genome map of DLP3. Scale in bp is shown on the outer periphery. Predicted functions are grouped by color: teal; moron, grey; hypothetical, light blue; DNA packaging, pink; tRNA, red; lysis, green; virion morphogenesis, dark blue; DNA replication and repair, purple; auxiliary metabolism, and yellow; regulatory.

Table 3-1: Genome annotations for DLP3 obtained from BLASTp and CD-Search data. Results below 0.01 were not used and the function was annotated as hypothetical.

CDS	Interval	Length (AA)	Putative function	Species	Coverage (%)	E-value	Identity (%)	Accession
1	48 - 1856	602	portal protein	DLP5[a]	98	0	83	ATS92275.1
2	1856 - 3193	445	ParB-like nuclease domain protein	DLP5	100	0	75	ATS92281.1
3	3249 - 3854	201	serine protease	DLP5	98	1.00E-113	80	ATS92325.1
4	3847 - 4017	56	hypothetical protein gp_005	DLP5	100	3.00E-24	77	ATS92409.1
5	4090 - 5307	405	hypothetical protein gp_006	DLP5	100	0	79	ATS92283.1
6	5340 - 6272	310	major capsid protein	DLP5	100	0	93	ATS92299.1
7	6358 - 6624	88	hypothetical protein gp_008	DLP5	100	3.00E-33	65	ATS92383.1
8	6693 - 7436	247	ribonuclease E	DLP5	100	3.00E-113	73	ATS92315.1
9	7457 - 8068	203	hypothetical protein gp_010	DLP5	100	2.00E-121	82	ATS92324.1
10	8068 - 8493	141	tail protein	DLP5	99	6.00E-71	77	ATS92347.1
11	8490 - 8936	148	hypothetical protein gp_012	DLP5	99	3.00E-82	76	ATS92342.1
12	9012 - 9305	97	hypothetical protein gp_013	DLP5	100	5.00E-50	85	ATS92376.1
13	9316 - 9693	125	hypothetical protein gp_014	DLP5	100	4.00E-48	67	ATS92360.1
14	9693 - 10448	251	hypothetical protein gp_015	DLP5	100	2.00E-152	82	ATS92314.1
15	10465 - 10959	164	hypothetical protein gp_016	DLP5	100	9.00E-83	81	ATS92333.1
16	11016 - 11144	42	hypothetical protein gp_017	DLP5	95	2.00E-15	83	ATS92415.1
17	11151 - 14915	1254	tape measure protein	DLP5	99	0	83	ATS92270.1
18	14917 - 16500	527	hypothetical protein gp_019	DLP5	100	0	79	ATS92277.1
19	16500 - 17474	324	hypothetical protein gp_020	DLP5	100	0	83	ATS92295.1
20	17474 - 19153	559	hypothetical protein gp_021	DLP5	100	0	74	ATS92276.1
21	19150 - 19968	272	minor tail protein	DLP5	100	0	90	ATS92307.1
22	19968 - 20255	95	hypothetical protein gp_023	DLP5	100	5.00E-62	98	ATS92377.1
23	20252 - 20455	67	hypothetical protein gp_024	DLP5	100	1.00E-38	93	ATS92399.1

24	20445 - 22925	826	tail protein	DLP5	100	0	86	ATS92271.1
25	22925 - 23923	332	tail assembly protein	DLP5	100	0	86	ATS92294.1
26	23926 - 24102	58	tail assembly protein	DLP5	100	3.00E-30	88	ATS92406.1
27	24110 - 25081	323	tail assembly protein	DLP5	100	2.00E-169	70	ATS92296.1
28	25085 - 26107	340	tail protein	Salvo[b]	100	3.00E-76	43	AHB12239.1
29	26183 - 26611	142	hypothetical protein gp_030	DLP5	100	1.00E-91	93	ATS92345.1
30	26611 - 27096	161	hypothetical protein gp_031	DLP5	100	6.00E-87	91	ATS92334.1
31	27096 - 27653	185	lysozyme	DLP5	100	2.00E-120	89	ATS92330.1
32	27655 - 28059	134	hypothetical protein gp_033	DLP5	100	3.00E-80	88	ATS92352.1
33	28110 - 28280	56	hypothetical protein gp_034	DLP5	78	1.00E-11	70	ATS92408.1
34	28329 - 28715	128	DUF2500 containing protein	DLP5	100	3.00E-77	89	ATS92350.1
35	28687 - 28974	95	hypothetical protein gp_036	DLP5	93	2.00E-57	96	ATS92380.1
36	29077 - 29271	64	hypothetical protein					
37	29275 - 29574	99	hypothetical protein gp_041	DLP5	95	2.00E-58	95	ATS92374.1
38	29634 - 30791	385	hypothetical protein gp_042	DLP5	100	0	78	ATS92287.1
39	30898 - 31506	202	hypothetical protein gp_043	DLP5	87	2.00E-82	70	ATS92323.1
40	31499 - 31951	150	phosphoglycerate kinase	DLP5	100	2.00E-103	96	ATS92341.1
41	31951 - 32136	61	hypothetical protein gp_046	DLP5	100	2.00E-25	79	ATS92401.1
42	32197 - 33141	314	hypothetical protein gp_047	DLP5	100	3.00E-138	85	ATS92300.1
43	33255 - 34856	533	hypothetical protein gp_048	DLP5	98	2.00E-156	50	ATS92279.1
44	34856 - 35137	93	hypothetical protein gp_049	DLP5	100	5.00E-35	64	ATS92378.1
45	35139 - 36056	305	hypothetical protein gp_050	DLP5	100	0	90	ATS92301.1
46	36060 - 36251	63	hypothetical protein	BCC[c]	98	5.00E-08	46	WP_046196969.1
47	36389 - 36571	60	hypothetical protein gp_052	DLP5	85	1.00E-19	78	ATS92407.1
48	36568 - 37128	186	hypothetical protein gp_053	DLP5	100	5.00E-128	93	ATS92329.1
49	37125 - 38414	429	helicase	DLP5	100	0	90	ATS92282.1
50	38481 - 38981	166	hypothetical protein gp_055	DLP5	100	2.00E-86	80	ATS92331.1

51	38974 - 39957	327	primase	DLP5	100	0	77	ATS92297.1
52	40036 - 40812	258	hypothetical protein gp_057	DLP5	100	3.00E-157	89	ATS92312.1
53	40860 - 41075	71	hypothetical protein gp_058	DLP5	100	4.00E-41	92	ATS92395.1
54	41072 - 42430	452	superfamily II DNA or RNA helicase	DLP5	100	0	90	ATS92280.1
55	42427 - 42828	133	transcriptional regulator	DLP5	100	6.00E-68	74	ATS92353.1
56	42831 - 43208	125	hypothetical protein gp_061	DLP5	100	1.00E-62	81	ATS92362.1
57	43208 - 44425	405	RecA	DLP5	100	0	90	ATS92285.1
58	44425 - 44799	124	hypothetical protein gp_063	DLP5	100	9.00E-83	94	ATS92359.1
59	44799 - 45680	293	hypothetical protein gp_064	DLP5	100	0	87	ATS92304.1
60	45677 - 46174	165	RuvC	DLP5	99	5.00E-111	93	ATS92332.1
61	46152 - 46766	204	hypothetical protein gp_066	DLP5	100	5.00E-108	78	ATS92321.1
62	46848 - 47657	269	hypothetical protein gp_067	DLP5	96	2.00E-153	83	ATS92308.1
63	47840 - 47947	35	hypothetical protein gp_069	DLP5	100	5.00E-13	94	ATS92417.1
64	48028 - 48504	158	hypothetical protein gp_070	DLP5	100	3.00E-96	91	ATS92336.1
65	48839 - 49174	111	hypothetical protein gp_071	DLP5	100	8.00E-60	82	ATS92357.1
66	50468 - 51409	313	hypothetical protein gp_073	DLP5	100	6.00E-127	61	ATS92313.1
67	51397 - 51585	62	hypothetical protein					
68	51582 - 52145	187	hypothetical protein gp_074	DLP5	100	2.00E-124	91	ATS92328.1
69	52165 - 52959	264	SPFH domain-containing protein	DLP5	100	0	96	ATS92309.1
70	53058 - 53210	50	hypothetical proteins					
71	53207 - 53653	148	hypothetical protein gp_077	DLP5	100	3.00E-87	87	ATS92343.1
72	53650 - 54000	116	hypothetical protein gp_078	DLP5	100	7.00E-74	90	ATS92365.1
73	53993 - 55207	404	hypothetical protein gp_079	DLP5	100	7.00E-96	44	ATS92288.1
74	55279 - 55548	113	hypothetical protein gp_081	DLP5	100	1.00E-65	84	ATS92368.1
75	55558 - 55899	89	hypothetical protein gp_080	DLP5	98	8.00E-37	66	ATS92369.1
76	55887 - 56930	347	UDP-glucose 4-epimerase	DLP5	100	0	83	ATS92289.1
77	56962 - 57285	107	hypothetical protein gp_083	DLP5	99	9.00E-56	78	ATS92370.1

78	57355 - 58425	356	hypothetical protein gp_084	DLP5	100	0	73	ATS92290.1
79	58418 - 58939	173	hypothetical protein gp_087	DLP5	45	6.00E-35	73	ATS92388.1
80	58936 - 59283	115	hypothetical protein gp_086	DLP5	86	1.00E-36	64	ATS92361.1
81	59280 - 59456	58	hypothetical protein gp_088	DLP5	100	4.00E-17	62	ATS92405.1
82	59465 - 59737	90	hypothetical protein BV378_14040	*Nostoc* sp. RF31Y	97	8.00E-19	50	OUL25853.1
83	59737 - 59862	41	hypothetical protein					
84	59859 - 60128	89	hypothetical protein					
85	60132 - 60314	60	hypothetical protein gp_089	DLP5	100	2.00E-05	53	ATS92403.1
86	60318 - 60737	139	hypothetical protein gp_090	DLP5	84	3.00E-64	79	ATS92363.1
87	60734 - 61210	158	hypothetical protein gp_091	DLP5	98	7.00E-60	62	ATS92339.1
88	61207 - 61458	83	hypothetical protein gp_092	DLP5	100	1.00E-22	46	ATS92382.1
89	61455 - 61709	84	hypothetical protein					
90	61709 - 62359	216	hypothetical protein gp_093	DLP5	100	3.00E-100	65	ATS92319.1
91	62362 - 62649	95	hypothetical protein gp_094	DLP5	96	4.00E-35	66	ATS92381.1
92	62646 - 62885	79	hypothetical protein gp_095	DLP5	100	4.00E-31	70	ATS92392.1
93	63161 - 63688	175	rhomboid membrane protein	DLP5	100	9.00E-96	78	ATS92322.1
94	63746 - 64342	198	hypothetical protein gp_097	DLP5	98	8.00E-108	82	ATS92326.1
95	64339 - 64593	84	hypothetical protein gp_098	DLP5	96	3.00E-31	68	ATS92385.1
96	64670 - 65308	212	PIG-L family deacetylase	DLP5	100	6.00E-114	75	ATS92320.1
97	65311 - 66252	313	WcaG	DLP5	100	0	95	ATS92298.1
98	66252 - 67298	348	WecE	DLP5	100	0	87	ATS92291.1
99	67295 - 67975	226	methyltransferase	DLP5	100	1.00E-139	82	ATS92317.1
100	68043 - 68228	61	hypothetical protein gp_103	DLP5	98	5.00E-20	62	ATS92400.1
101	68225 - 68647	140	hypothetical protein gp_104	DLP5	97	8.00E-56	68	ATS92346.1
102	68647 - 70095	482	n-acetyl-alpha-d-glucosaminyl l-malate synthase	DLP5	100	0	93	ATS92278.1

103	70092 - 70403	103	hypothetical protein gp_106	DLP5	100	1.00E-36	59	ATS92372.1
104	70375 - 71229	284	ParBc	DLP5	99	0	89	ATS92305.1
105	71229 - 71879	216	hypothetical protein gp_108	DLP5	99	5.00E-125	85	ATS92318.1
106	71842 - 73755	637	terminase large subunit	DLP5	100	0	91	ATS92274.1
107	73767 - 74660	297	hypothetical protein gp_110	DLP5	100	0	95	ATS92303.1
108	74657 - 75067	136	DUF3310 containing protein	DLP5	100	1.00E-61	71	ATS92351.1
109	75129 - 75371	80	hypothetical protein gp_112	DLP5	100	8.00E-37	74	ATS92390.1
110	75373 - 75765	130	hypothetical protein gp_113	DLP5	92	7.00E-12	34	ATS92355.1
111	75841 - 75939	32	hypothetical protein gp_114	DLP5	100	7.00E-09	81	ATS92418.1
112	75936 - 76253	105	hypothetical protein gp_115	DLP5	96	6.00E-54	81	ATS92371.1
113	76250 - 76504	84	hypothetical protein gp_116	DLP5	100	2.00E-42	79	ATS92387.1
114	76494 - 76859	121	hypothetical protein gp_117	DLP5	98	6.00E-59	80	ATS92358.1
115	76859 - 77011	50	hypothetical protein gp_118	DLP5	100	1.00E-22	88	ATS92412.1
116	77087 - 77335	82	hypothetical protein gp_119	DLP5	100	2.00E-43	82	ATS92389.1
117	77389 - 78528	379	hypothetical protein gp_120	DLP5	70	2.00E-143	94	ATS92286.1
118	78528 - 78878	116	hypothetical protein gp_121	DLP5	99	4.00E-71	90	ATS92364.1
119	78904 - 79365	153	hypothetical protein gp_122	DLP5	100	7.00E-78	75	ATS92340.1
120	79331 - 79561	76	hypothetical protein gp_123	DLP5	100	3.00E-33	76	ATS92393.1
121	79558 - 79731	57	hypothetical protein gp_124	DLP5	96	6.00E-23	78	ATS92404.1
122	79804 - 79920	38	hypothetical protein					
123	79920 - 80228	102	hypothetical protein gp_125	DLP5	100	3.00E-59	87	ATS92373.1
124	80254 - 80988	244	hypothetical protein gp_126	DLP5	100	8.00E-153	83	ATS92316.1
125	80985 - 81350	121	hypothetical protein gp_127	DLP5	96	1.00E-64	79	ATS92356.1
126	81347 - 81523	58	hypothetical protein					
127	81516 - 81698	60	hypothetical protein gp_128	DLP5	96	2.00E-29	86	ATS92402.1
128	81698 - 82180	160	DUF1643 containing protein	DLP5	100	2.00E-88	80	ATS92337.1
129	82180 - 82515	111	hypothetical protein gp_130	DLP5	90	2.00E-44	83	ATS92367.1

130	82589 - 83485	298	DNA ligase	DLP5	100	0	92	ATS92302.1
131	83482 - 83895	137	hypothetical protein gp_133	DLP5	100	4.00E-65	74	ATS92349.1
132	83895 - 84197	100	hypothetical protein gp_134	DLP5	97	2.00E-47	77	ATS92375.1
133	84248 - 85459	403	hypothetical protein gp_135	DLP5	100	0	81	ATS92284.1
134	85456 - 86028	190	hypothetical protein gp_136	DLP5	100	3.00E-119	85	ATS92327.1
135	86122 - 88065	647	pyruvate phosphate dikinase	DLP5	100	0	84	ATS92273.1
136	88141 - 88527	128	hypothetical protein gp_138	DLP5	98	7.00E-76	90	ATS92348.1
137	88527 - 88868	113	hypothetical protein gp_139	DLP5	100	5.00E-63	86	ATS92366.1
138	88879 - 89238	119	transcriptional repressor	DLP5	95	3.00E-71	88	ATS92354.1
139	89228 - 89698	156	tyrosine phosphatase family protein	DLP5	99	5.00E-93	86	ATS92338.1
140	89740 - 90759	339	hypothetical protein gp_142	DLP5	98	0	77	ATS92292.1
141	90756 - 91532	258	thymidylate synthase	DLP5	98	6.00E-121	72	ATS92311.1
142	91618 - 92493	291	hypothetical protein gp_144	DLP5	97	1.00E-124	65	ATS92306.1
143	92483 - 92737	84	hypothetical protein gp_145	DLP5	97	3.00E-35	71	ATS92386.1
144	92734 - 93231	165	hypothetical protein gp_146	DLP5	96	1.00E-83	77	ATS92335.1
145	93260 - 95482	740	DNA polymerase I	DLP5	100	0	85	ATS92272.1
146	95794 - 96030	78	hypothetical protein					
147	96134 - 96415	93	hypothetical protein gp_149	DLP5	96	6.00E-52	87	ATS92379.1
148	96412 - 96741	109	hypothetical protein gp_150	DLP5	100	2.00E-61	85	ATS92310.1

[a] *Stenotrophomonas* phage DLP5, [b] *Xylella* phage Salvo, and [c]*Burkholderia cenocepacia* complex

Table 3-2: The conserved domains found in the 148 DLP3 gene products.

Gp	Hit type	PSSM-ID	Interval	E-Value	Accession	Short name	Superfamily
1	superfamily	327517	29 - 471	2.89E-49	cl19194	Phage_portal superfamily	-
2	specific	214678	359 - 441	8.46E-06	smart00470	ParB	cl02129
3	superfamily	317012	5 - 114	7.39E-23	cl24270	Peptidase_S78_2 superfamily	-

6	superfamily	331903	20 - 307	5.06E-12	cl27082	Phage_capsid superfamily	-
10	superfamily	321796	6 - 137	7.98E-15	cl02089	Phage_tail_S superfamily	-
17	superfamily	331332	525 - 930	5.90E-15	cl26511	Neuromodulin_N superfamily	-
17	superfamily	333387	163 - 230	3.22E-03	cl28567	HI1514 superfamily	-
19	superfamily	316645	14 - 69	9.95E-03	cl16644	DUF4302 superfamily	-
21	superfamily	312753	188 - 264	8.54E-16	cl10710	Phage_BR0599 superfamily	-
21	superfamily	331404	18 - 261	5.62E-12	cl26583	DUF2163 superfamily	-
24	specific	316107	207 - 368	2.18E-13	pfam13550	Phage-tail_3	cl26145
28	superfamily	328935	62 - 184	1.13E-05	cl22861	LamG superfamily	-
31	superfamily	331815	6 - 185	4.73E-34	cl26994	Glyco_hydro_108 superfamily	-
38	superfamily	332389	92 - 371	4.11E-42	cl27568	TIP49 superfamily	-
41	superfamily	332243	13 - 46	5.11E-03	cl27422	SecD superfamily	-
49	superfamily	333705	167 - 405	5.02E-22	cl28885	RecA-like_NTPases superfamily	-
51	superfamily	331610	31 - 322	7.81E-12	cl26789	Toprim_N superfamily	-
54	superfamily	331760	51 - 451	4.03E-38	cl26939	DEXDc superfamily	-
55	superfamily	322007	53 - 98	3.22E-04	cl02600	HTH_MerR-SF superfamily	-
57	superfamily	333705	77 - 272	2.57E-48	cl28885	RecA-like_NTPases superfamily	-
60	superfamily	328743	1 - 140	6.73E-14	cl21482	RuvC_resolvase superfamily	-
69	specific	307341	23 - 207	2.15E-17	pfam01145	Band_7	cl19107
76	specific	224012	1 - 336	1.19E-76	COG1087	GalE	cl21454

78	superfamily	330522	59 - 126	2.19E-04	cl25701	RuvB_N superfamily	-
93	superfamily	328780	29 - 159	3.68E-14	cl21536	Rhomboid superfamily	-
96	specific	308281	5 - 122	3.46E-11	pfam02585	PIG-L	cl00929
97	specific	223528	1 - 297	7.51E-40	COG0451	WcaG	cl25660
98	superfamily	327488	20 - 340	5.46E-56	cl18945	AAT_I superfamily	-
99	superfamily	327401	30 - 135	1.43E-04	cl17173	AdoMet_MTases superfamily	-
102	specific	223515	1 - 323	4.33E-03	COG0438	RfaB	cl28208
104	specific	214678	18 - 109	7.28E-17	smart00470	ParB	cl02129
104	specific	224392	34 - 187	5.36E-08	COG1475	Spo0J	cl26722
108	superfamily	314594	17 - 68	1.07E-11	cl13237	DUF3310 superfamily	-
128	specific	311648	12 - 146	4.16E-44	pfam07799	DUF1643	cl01787
130	superfamily	325160	25 - 186	2.14E-25	cl12015	Adenylation_DNA_ligase_like superfamily	-
130	superfamily	330238	112 - 270	1.28E-08	cl25417	CDC9 superfamily	-
133	superfamily	332389	206 - 341	8.78E-33	cl27568	TIP49 superfamily	-
133	superfamily	332204	43 - 150	8.52E-03	cl27383	ERCC4 superfamily	-
135	superfamily	331842	16 - 471	0.00E+00	cl27021	PtsP superfamily	-
139	superfamily	330819	39 - 125	1.50E-10	cl25998	CDC14 superfamily	-
141	superfamily	330819	85 - 189	1.80E-05	cl25998	CDC14 superfamily	-
145	superfamily	322025	107 - 710	7.76E-64	cl02626	DNA_pol_A superfamily	-

Phyre2 analysis of these proteins only suggested a putative function for gp14, which showed structural identity to the major tropism determinant p1 (MTD-P1) of *Bordetella* phage BPP-1 (38.7 % confidence, 25 % identity). The MTD-P1 protein in BPP-1 is responsible for phage receptor binding[232] and uses specific variable residues for target recognition similar to an antibody[233]. Another abundant protein identified is gp24, which is predicted to be a structural component of the DLP3 tail with a phage-tail_3 family (PF13550) conserved domain. The tape measure protein (gp17), which is responsible for setting the tail length[234], is the largest protein identified and the sixth most abundant, though the ~130 kDa band was faint after fully de-staining the gel (Figure 3-3). The remaining DLP3 proteins identified and ordered by relative abundance are gene products 19, 97, 27, 61, 20, 28, 2, 21, 25, 9, 133, 12, 57, 5 and 3. Some of the proteins identified are surprising, such as gp2 (Par-B like nuclease domain protein) and gp5 (RecA), but based on the results obtained by screening the *Stenotrophomonas* protein database, it is evident that there were proteins present from the bacterial cell; thus, there may also be DLP3 proteins present that are not structural in function (Table 3-3).

By screening the *Stenotrophomonas* database, numerous proteins were identified (Table 3-4). The D1571 host strain proteins are not in the UniProt database; but, screening the peptide sequences against predicted protein sequences from a D1571 scaffold has enabled the identification of these proteins in D1571. There are many TonB-dependent receptor proteins identified with the database screen (Table 3-4), but only two large proteins in higher abundance will be discussed further as they are visible on the PAGE gel used for the mass spectrometry analysis (Figure 3-3). The Oar protein of *S. maltophilia* strain K279a (SMLT_RS07785) was identified (84 PSMs, 118 kDa) and a conserved domain search of this protein revealed a carboxypeptidase regulatory-like domain (pfam13620, 5.75e-20) and the outer membrane receptor protein domain CirA (COG1629, 2.61e-08). When screening the peptide sequences from the Oar protein against putative D1571 proteins, a protein was identified with 100% amino acid identity.

The next most abundant bacterial protein identified based on the number of PSMs (94 PSMs, 105.6 kDa) is labeled as a TonB-dependent receptor (B9Y71_14085) and a CD-Search of this protein identified a porin superfamily domain (cl21487, 6.86e-45), which was also found within a D1571 protein identified (cl21487, 1.46e-45). The D1571 protein shares 99.5 % identity with the TonB-dependent receptor with five amino acid substitutions identified. The abundance

of these proteins in the results is surprising, but because the phage lysate was not ultra-purified with a CsCl gradient, it is not entirely unexpected. It could be argued that one of the outer membrane proteins identified is the primary receptor for DLP3 and was enriched in the sample by maintaining a stable association with the DLP3 receptor binding protein, but further investigation is required to substantiate this claim. Preliminary results obtained from a graduate student in the Dennis lab, Jaclyn McCutcheon, has suggested DLP3 uses the type IV pili as the secondary receptor, though DLP3 has not been shown to use the porin PilQ (~77 kDa) for DNA injection.

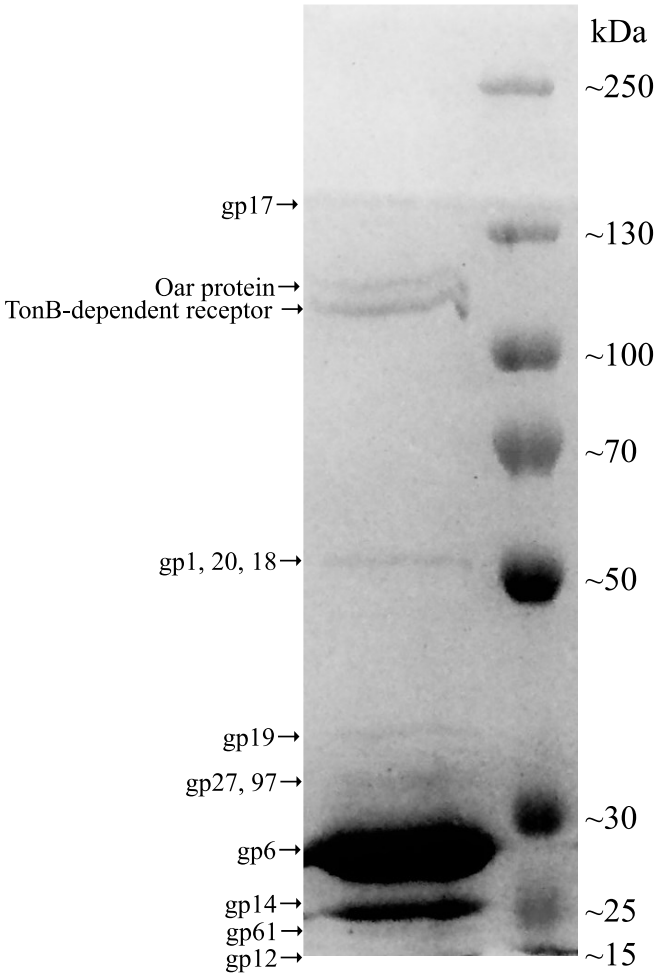

Figure 3-3: SDS-PAGE gel of DLP3 ghost particles (L) compared to a PageRuler Plus Pre-stained Protein Ladder (Thermo Scientific). Approximate masses are labelled on the right of the gel in kDa. Protein bands labelled using mass spectrometry results from SEQUEST scan of DLP3 and *Stenotrophomonas* protein databases. Stacking and resolving portions of the gel were 4 and 7.5% respectively. The gel was stained with Coomassie R-250 and imaged with an iPhone 5S

Table 3-3: Mass spectrometry protein results with the DLP3 protein database. Results are ordered by score.

Gp	Function	Score	Coverage (%)	Unique Peptides	Peptides	PSMs	AAs	MW [kDa]	calc. pI
6	major capsid protein	1635.14	91.94	27	27	2709	310	33.7	5.76
14	hypothetical protein gp_015	189.02	56.57	12	12	282	251	27.4	4.92
1	portal protein	95.51	51.24	35	36	339	566	62.7	5.41
18	hypothetical protein gp_019	54.48	19.54	10	10	123	527	58.2	6.27
17	tape measure protein	45.75	21.85	28	28	78	1254	131.9	6.39
27	tail assembly protein	33.34	28.48	7	7	46	323	35.4	7.05
19	DUF4302 family protein	30.84	27.47	7	7	69	324	35.8	6.61
24	tail protein	23.39	26.51	18	18	106	826	90.1	5.08
97	WcaG	22.78	30.99	10	10	52	313	34.8	5.99
21	minor tail protein	11.39	28.68	9	9	19	272	30.5	7.33
2	ParB-like nuclease domain protein	11.28	20.45	8	8	19	445	49.3	6.96
9	hypothetical protein gp_010	8.01	31.03	6	6	11	203	22.5	5.94
28	tail protein	7.74	24.71	10	10	24	340	37.2	8.73
20	hypothetical protein gp_021	7.56	14.85	9	9	30	559	63.5	5.25
61	hypothetical protein gp_066	6.45	27.45	7	7	31	204	22.3	9.00
25	tail assembly protein	4.15	23.80	6	6	15	332	35.5	5.15
12	hypothetical protein gp_013	2.37	78.35	4	4	7	97	9.6	9.47
3	serine protease	2.24	9.95	2	2	3	201	22.7	4.67
57	RecA	2.03	15.56	5	5	7	405	43.5	6.46
5	hypothetical protein gp_006	0.00	6.67	2	2	4	405	43.4	5.39
133	ATPase family protein	0.00	10.67	4	4	9	403	44.6	5.80

Table 3-4: Mass spectrometry protein results using the *Stenotrophomonas* protein database in UniProt. The results are organized by score.

Hit	Gene	Score	Coverage (%)	Proteins	Unique Peptides	Peptides	PSMs	AAs	MW [kDa]	calc. pI
TonB-dependent receptor	B9Y71_14085	127.60	23.45	3	20	20	94	985	105.6	6.42
Oar protein	BWP19_02175	115.44	22.70	3	22	22	84	1075	118.0	6.48
TonB-dependent receptor	AR275_14455	72.57	22.09	1	18	18	64	987	106.9	5.36
Putative TonB dependent receptor	Smlt3478[a]	69.96	14.94	1	12	12	43	944	100.4	5.67
TonB-dependent receptor	BWP19_04650	49.49	11.48	2	10	10	26	1028	111.9	6.46
Porin	B9Y71_15810	48.38	29.56	3	11	11	25	389	42.9	6.64
Uncharacterized protein	AR275_04505	41.26	38.25	1	11	11	27	366	39.5	5.00
TonB-dependent receptor	BWP19_01190	39.51	18.17	4	15	15	31	952	103.0	5.83
Ligand-gated channel	YH67_15830	36.92	15.29	11	10	10	19	811	87.5	5.59
TonB-dependent receptor	YH67_02440	32.07	11.49	20	10	10	22	879	94.4	6.62
Probable cytosol aminopeptidase	pepA	30.69	12.40	11	6	6	17	492	51.2	5.52
Uncharacterized protein	B9Y76_18325	26.74	37.13	1	7	7	14	272	30.0	5.41
Dihydrolipoyl dehydrogenase	B9Y71_11945	23.96	14.85	4	7	7	17	478	50.4	6.70
Protein CyaE	B9Y71_15750	22.44	17.04	13	6	6	12	452	48.7	6.46
Membrane protein	ABW44_09025	19.18	14.89	1	5	5	11	450	47.2	6.55
50S ribosomal protein L2 (Fragment)	VM57_03345	16.67	23.74	5	4	4	10	198	21.3	9.99
DNA-directed RNA polymerase subunit	L681_20760[c]	16.51	8.99	28	7	7	10	790	88.9	8.97
Uncharacterized protein	B9Y57_19230	15.88	38.83	1	6	6	10	206	22.6	8.29

TonB-dependent receptor	VL23_08990	13.73	9.52	18	7	7	10	756	82.7	5.36
30S ribosomal protein S13	rpsM	13.36	31.36	3	3	3	6	118	13.4	11.46
TonB-dependent receptor	STRNTR1_0356	13.15	5.98	24	6	6	7	953	103.1	5.45
OmpW family protein	SmaCSM2_17180	12.97	20.57	21	3	3	12	209	22.1	8.76
TonB-dependent receptor	BWP19_04640	11.53	7.49	3	5	6	10	935	101.8	6.30
Esterase	AR275_13845	11.17	8.51	3	4	4	7	611	64.1	6.55
Cell surface protein	BWP19_07340	10.67	2.47	6	5	5	9	2393	223.5	4.59
Autotransporter-associated beta strand	A1OC_03319[b]	9.41	4.87	4	4	4	5	944	96.6	8.22
TonB-dependent receptor	AR275_02775	8.50	13.05	5	7	7	8	613	66.5	6.65
Porin	SmaCSM2_17905	8.17	11.08	4	4	4	4	379	41.9	7.08
Malic enzyme (Fragment)	AR275_14665	7.89	3.22	22	2	2	4	652	70.9	6.19
30S ribosomal protein S3	rpsC	7.46	23.24	5	4	4	5	241	27.2	10.15
TonB-dependent receptor	ABW44_04685	7.41	7.55	7	5	6	8	940	103.1	5.66
TonB-dependent receptor	B9Y56_09770	7.28	3.15	2	3	3	5	1047	114.4	5.57
Outer membrane receptor protein	BB780_13710	5.97	2.94	7	2	2	3	749	80.7	6.51
TonB-dependent receptor	BWP19_12960	5.50	2.96	19	3	3	5	912	100.4	5.60
Uncharacterized protein	CR919_19680	4.41	3.96	23	2	2	2	732	79.3	5.59
ATP synthase gamma chain	atpG	4.31	8.36	8	2	2	2	287	31.9	9.61

TldD protein (Fragment)	AR275_13300	4.18	4.52	9	2	2	2	442	47.4	8.07
Ribonucleoside-diphosphate reductase	RRM1[a]	4.07	5.68	14	3	3	3	775	86.7	6.57
Autotransporter outer membrane beta-barrel domain-containing protein	B9Y56_12910	3.83	2.65	1	2	2	2	791	82.4	6.06
Oar protein	BWP19_04665	3.79	2.68	2	2	2	2	1007	109.5	5.81
Uncharacterized protein	ARC23_12465	3.73	3.05	1	3	3	3	985	108.6	5.47
TonB-dependent receptor	AR275_05645	3.73	5.08	8	3	3	3	768	82.6	6.57
Flagellin	flic7[d]	3.67	6.67	8	2	2	3	390	40.0	5.40
TonB-dependent receptor (Fragment)	ARC78_03340	3.63	3.52	31	2	2	3	739	79.3	5.22
30S ribosomal protein S2	rpsB	3.61	10.49	9	3	3	4	267	29.4	8.51
TonB-dependent receptor	ABW44_01270	3.60	4.59	2	4	4	5	806	88.6	6.46
TonB-dependent receptor	YH67_13890	3.51	5.15	27	4	4	4	776	84.5	6.10
Porin	B9Y61_10620	3.44	4.00	22	2	2	2	475	50.9	6.09
Peptidase M20	YH67_02165	3.37	7.29	22	3	3	3	549	59.7	6.43
Uncharacterized protein	u35[d]	2.52	6.69	2	3	3	3	658	71.5	5.50
ATP synthase subunit alpha	atpA	1.89	3.88	1	2	2	2	515	55.3	5.62
Uncharacterized protein (Fragment)	BWP19_21350	1.80	9.52	9	2	2	2	210	23.0	4.65
Histidine ammonia-lyase	hutH	1.79	5.85	7	3	3	3	513	53.5	5.69
OmpA-related protein	SmaCSM2_14405	1.71	3.73	15	3	3	4	1018	110.4	6.37

[a]K279a, [b]Ab55555, [c]MF89, [d]RA8

Lysogenic conversion of D1571 by temperate phage DLP3

Stable lysogens of D1571::DLP3 have been isolated, and due to the presence of ParB domains in two of the DLP3 proteins, DLP3 may lysogenize its host as a phagemid. The closely related phage DLP5 was also found to encode two proteins with ParB domains and has successfully been isolated as a plasmid from the lysogenized strain D1614 (Chapter 5). DLP5 was shown to cause lysogenic conversion of its host; therefore, further investigation into the D1571::DLP3 lysogen was warranted.

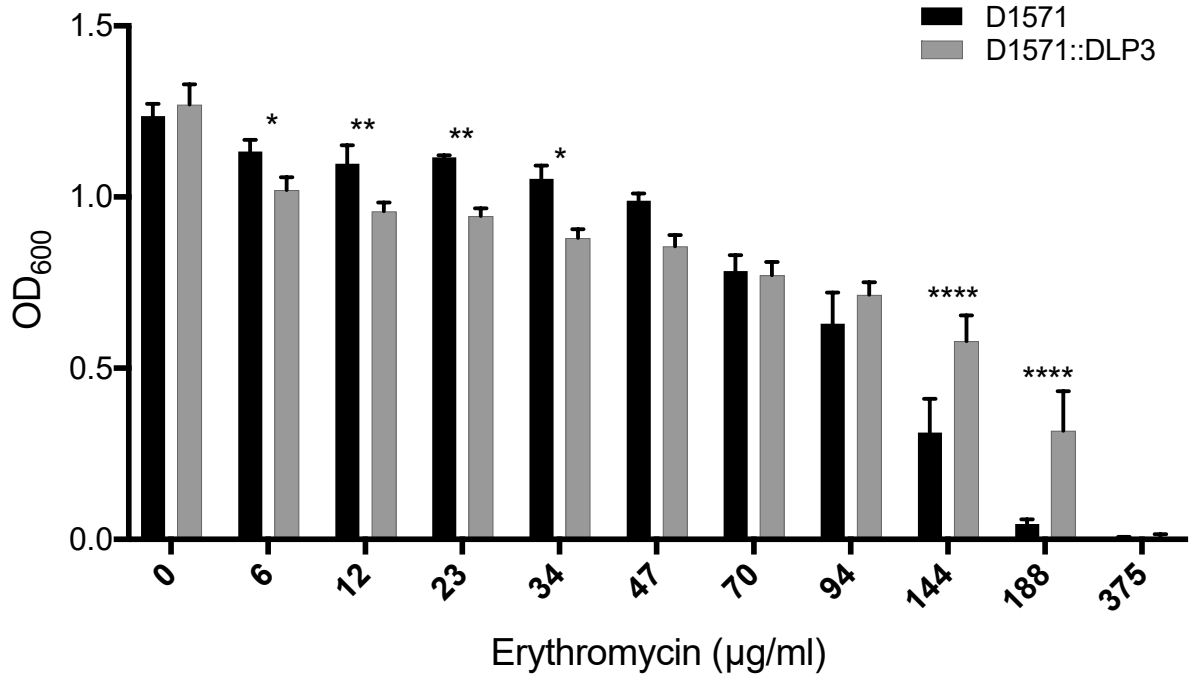

Figure 3-4: Erythromycin resistance of D1571::DLP3 lysogen increases compared to wild type control D1571. Minimum inhibitory concentration assay was completed in biological and mechanical triplicates. Two-way ANOVA with Sidak's multiple comparisons test was performed on the MIC data. Statistical significance is represented as: ****, $P < 0.0001$; **, $P < 0.01$; and *, $P < 0.05$.

Identification of erythromycin resistance factor Erm(45)

Annotation of the DLP3 genome revealed the presence of a methyltransferase domain in the gene product encoded by DLP3_099. A specific domain identified is the AdoMet_MTases superfamily comprised of class I S-adenosylmethionine-dependent methyltransferases. The class I family is the largest and most diverse, with members targeting a range of substrate specificities

such as small molecules, lipids, nucleic acids, and different target atoms for methylation[235]. Members of this family are known to maintain structural similarity even when the amino acid sequence diverges to as little as 10 %[236]. Further investigation into gp69 with HHblits revealed high sequence identity to Erm(45) of *Staphylococcus fleurettii* (HHblits: 100 % probability, 2.4e-104 e-value). Researchers identified Erm(45) as the enzyme responsible for increased erythromycin resistance in some strains of *S. fleuretti*[237]. The *erm* gene (erythromycin ribosome methylase) encodes a methylase enzyme which provides macrolide resistance by methylating the erythromycin target-site on the ribosome[238]. Together, these findings suggested the DLP3 encoded methyltransferase is an erythromycin resistance protein similar to Erm(45) and could cause an increase in erythromycin resistance of the host during lysogeny.

To test the functionality of the DLP3 Erm protein, minimum inhibitory concentration (MIC) assays of the wild type D1571 and lysogen D1571::DLP3 were completed. The results show a statistically significant difference in erythromycin resistance with DLP3 lysogeny over a wide range of concentrations tested (Figure 3-4). There are differences noted between the wild type and lysogen strains at lower concentrations of erythromycin, with wild type having a statistically significant increase in optical density at lower concentrations (6 and 34 µg/ml; $P < 0.05$, 12 and 23 µg/ml; $P < 0.01$). The cause of the decreased resistance in the lysogen noted at lower concentrations of erythromycin is currently unknown. This trend is reversed at higher concentrations of erythromycin, with the lysogen having significantly higher OD600 readings at 144 and 188 µg/ml erythromycin concentrations compared to wild type control ($P < 0.0001$). The increased resistance noted at higher erythromycin concentrations shows the lysogenic conversion of the D1571 host by DLP3 and confirms the functionality of the predicted Erm protein.

Differing growth characteristics with DLP3 lysogenization

While working with the D1571::DLP3 lysogen and wild type D1571 for the erythromycin resistance assay, it was evident that the lysogen has a faster growth rate when compared to the wild type strain. This difference in growth rate was also observed when growing single colony isolates for the lysogen and wild type on an LB agar plate, with the lysogen having larger single colonies after a 16 h incubation at 30 ˚C compared to the wild type D1571 (data not shown). To investigate this phenotypic difference further, a growth curve was conducted in mechanical and biological triplicates for the lysogen and wild type over an eight-hour period. The plotted growth curve clearly shows the lysogenized strain exhibiting a faster growth rate than the wild type

strain (Figure 3-5). This observation was confirmed by converting the OD600 data into growth rate (μ) and plotting the resulting data (Figure 3-6). A two-way ANOVA with multiple comparisons revealed statistically significant (P <0.0001) differences between zero to two and two to four-hour time intervals for the lysogen and wild type D1571. The increased growth rate is only observed during the lag and early exponential phase of growth and disappears in the four to six and six to eight-hour growth rate intervals. The cause of the growth rate differences observed with DLP3 lysogenization is currently unknown, though DLP3 does encode many hypothetical and moron genes which may be expressed to produce this phenotype.

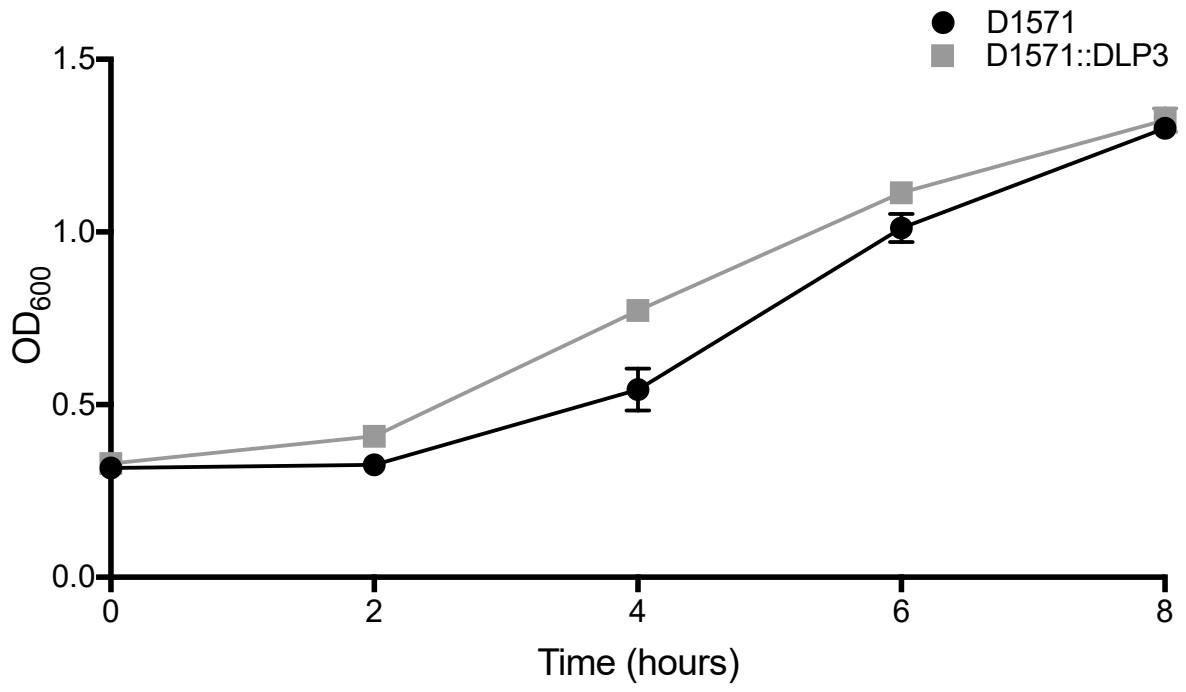

Figure 3-5: Growth curve analysis of D1571 wild type and D1571::DLP3 lysogen grown in LB broth over eight hours. Results from biological and mechanical triplicate experiments were averaged and the mean plotted with the standard deviation represented as error bars.

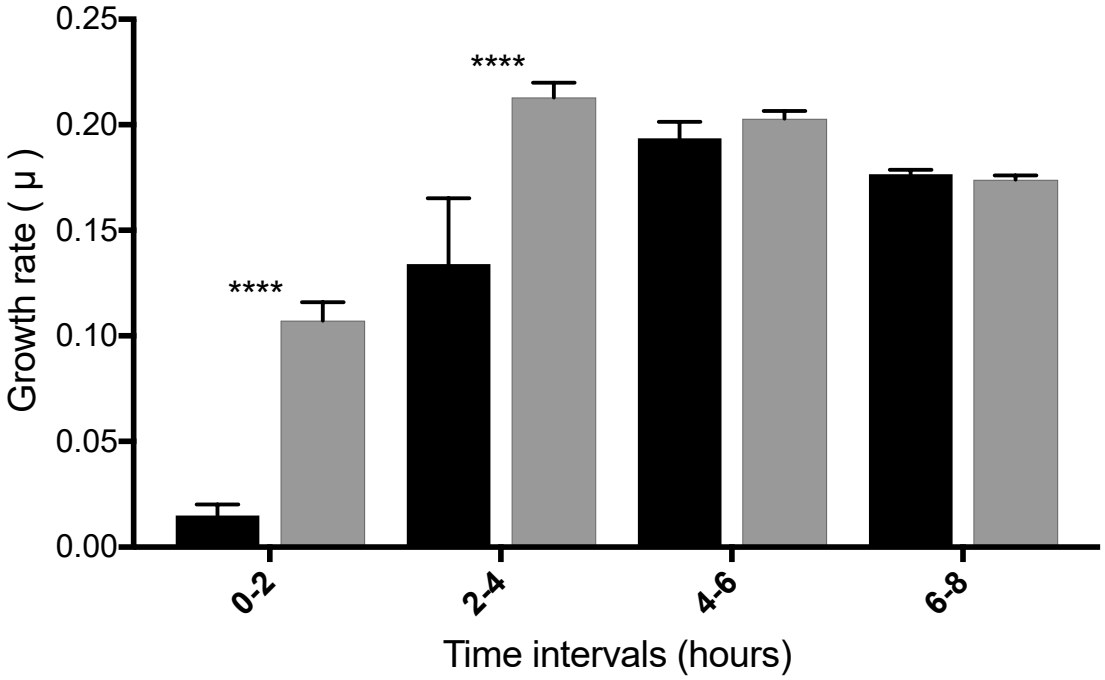

Figure 3-6: Growth rate of wild type *S. maltophilia* D1571 compared to lysogen D1571::DLP3. OD600 data from biological and mechanical triplicate experiments was converted to the growth rate and the mean plotted with the error bars representing the standard deviation. Two-way ANOVA with multiple comparisons indicates statistical significance with a *P* value of <0.0001.

CONCLUSIONS

The novel temperate phage DLP3 isolated from an Edmonton soil sample enabled the identification of a second member of the newly established *Delepquintavirus* genus. DLP3 has a genome size of 96,852 bp and a GC content of 58.4 %, which is significantly lower than the host strain D1571 which has a GC content of 66.9 %. DLP3 encodes two proteins with ParB conserved domains enabling the stable lysogenization of the host strain D1571. Lysogenization by DLP3 leads to a statistically significant growth rate increase during the lag and early exponential phase of growth for the host when compared to the wild type strain. DLP3 also encodes a functional Erm protein, allowing for the lysogenic conversion of the host D1571 strain which is observed though a statistically significant increased resistance to erythromycin at 144 and 188 µg/ml concentrations. All of the *Stenotrophomonas maltophlia* temperate phages identified and characterized for to date feature various moron genes which encode virulence factors or antibiotic resistance proteins which could lead to the lysogenic conversion of their host

[154,156,159,161,166]. These results highlight the role of phages in the evolution of more antibiotic-resistant or virulent *S. maltophilia* strains.

ACKNOWLEDGEMENTS

The authors would like to thank Arlene Oatway from the University of Alberta Department of Biological Sciences Advanced Microscopy Facility for assistance with electron microscopy. The authors thank the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR, Vancouver, BC) and The Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services, for gifts of the *S. maltophilia* strains used in this study.

Chapter 4

Lysogenic conversion of *Stenotrophomonas maltophilia* host D1585 by temperate siphovirus DLP4

Portions of this chapter will be submitted as:

Peters, D. L., McCutcheon, J. G., and J. J. Dennis. 2018. Lysogenic conversion of *Stenotrophomonas maltophilia* host D1585 by temperate siphovirus DLP4.

OBJECTIVES

The objectives of this project were fourfold: a) sequence and characterize the DLP4 genome, b) determine lifestyle, c) study functionality of moron genes and d) determine if lysogenic conversion is occurring in the *S. maltophilia* host D1585.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Five *S. maltophilia* and eight *P. aeruginosa* strains came from the Canadian *Burkholderia cepacia* complex Research and Referral Repository (Vancouver, BC). The *S. maltophilia* strain used to isolate DLP4 from the soil sample was D1585. An additional 22 *S. maltophilia* strains given to the lab from the Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services, for host range analysis. *Escherichia coli* BW25113 and *ybiA*770(del)::kan strains were used for swarming motility experiments and data from three biological and mechanical triplicates was used (Table 4-1). All strains were grown aerobically overnight at 30 °C on Lysogeny broth (LB or ½ LB) solid media or in LB or ½ LB broth with shaking at 225 RPM. Chloramphenicol at a final concentration of 35 µg/ml added when required.

Table 4-1: Bacterial strains and plasmids used in the expression studies.

Strain	Description	Source
Escherichia coli BW25113	Wild type control for Keio library	Baba, et al. [239]
Escherichia coli DH5α	Sub-clone host	Hanahan, et al. [240]
Stenotrophomonas maltophilia D1585	Wild type host for DLP4	CBCCRRR*
D1585::DLP4	*S. maltophilia* D1585 lysogen with DLP4 prophage	This study
pBBR	pBBR1MCS broad-host range cloning vector	Kovach, et al. [241]
pYbiA	pBBR1MCS carrying DLP4 *ybiA*, Cm^R	This study
pFolA	pBBR1MCS carrying DLP4 *folA*, Cm^R	This study

* Canadian *Burkholderia cepacia* complex Research and Referral Repository

Phage isolation, propagation, host range analysis, and electron microscopy

Bacteriophage DLP4 was isolated from a soil sample collected from Emily Murphy Park in Edmonton, Alberta, Canada using established protocols[167]. Briefly, the soil sample was incubated shaking at 30 °C in ½ LB broth with modified suspension medium (SM) and *S. maltophilia* D1585 liquid culture. The lysate was clarified by centrifugation, and the supernatant was filter-sterilized using a Millex-HA 0.45 µm syringe driven filter unit (Millipore). The lysate was mixed with strain D1585, plated in soft agar overlays, and incubated overnight at 30 °C. Single plaques were each isolated using a sterile Pasteur pipette and suspended in separate microcentrifuge tubes containing 500 µl SM with 20 µl chloroform and rocked one hour at room temperature.

DLP4 was propagated using soft agar overlays: 100 µl overnight culture and 100 µl phage stock incubated 20 min at room temperature, mixed with 3 ml 0.7 % ½ LB top agar, overlaid on a plate of ½ LB solid medium, and incubated at 30 °C overnight. High titer stocks were made by overlaying plates of near-confluent lysis with 3 ml SM and incubated >1 h at room temperature on a platform rocker. Centrifugation of the supernatant for 5 min at 10,000 x g

clarified the lysate which was then filter-sterilized using a Millex-HA 0.45 µm syringe-driven filter unit (Millipore) and stored at 4 °C.

Host range analysis was performed using 27 clinical *S. maltophilia* and 19 *P. aeruginosa* strains. Soft-agar overlays containing 100 µl liquid bacterial culture were allowed to solidify for 10 min at room temperature and spotted with 10 µl drops of DLP4 at multiple dilutions and assayed for clearing or plaque formation after overnight incubation at 30 °C.

For electron microscopy, phage stocks prepared as described above with the following modifications: ½ LB agarose plates and ½ LB soft agarose used for overlays, MilliQ-filtered water for phage recovery and passed through a 0.22 µm filter. A carbon-coated copper grid was incubated with lysate for 2 min and stained with 4 % uranyl acetate for 30 s. Transmission electron micrographs were captured using a Philips/FEI (Morgagni) transmission electron microscope with charge-coupled device camera at 80 kV (University of Alberta Department of Biological Sciences Advanced Microscopy Facility). The capsid diameter, length, and tail length were calculated using Microsoft Excel based on measurements from ten individual virions obtained using ImageJ software version 1.50i (NIH, Bethesda, MD).

Phage DNA isolation, RFLP analysis, and sequencing

DLP4 genomic DNA was isolated from bacteriophage lysate using the Wizard DNA purification system (Promega Corp.) with a modified protocol[168,169]. A NanoDrop ND-1000 spectrophotometer (Thermo Scientific) was used to determine the purity and concentration of eluted DNA. Restriction fragment length polymorphism analysis (RFLP) was performed using a panel of 36 FastDigest restriction enzymes (Thermo Scientific) and l µg of phage DNA. Reactions were incubated at 37 °C for 45 min and separated on a 0.8 % (wt/vol) agarose gel in 1x TAE (pH 8.0).

The DNA library was made by The Applied Genomics Core at the University of Alberta using a Nextera XT library prep kit and used for paired-end sequencing on a MiSeq (Illumina) platform using a MiSeq v2 reagent kit.

Bioinformatic analysis

Paired-end reads were assembled using SPAdes 3.8.0[215]. Open reading frames (ORFs) were identified using the GLIMMER plugin[172] for Geneious[217] using the Bacteria and Archaea setting as well as the GeneMarkS[174] program for phage. Conserved domain searches were

performed using CD-Search[175]. The contig was annotated using BLASTn and BLASTp (for full genomes and individual proteins, respectively)[176]. BLASTx and PHAST[242] were used to search for similar sequences in the GenBank database. Lysis protein analysis was performed using TMHMM for transmembrane region identification[179] and LipoP 1.0 for the prediction of lipoproteins[191]. Protein structure prediction was accomplished using I-TASSER[221]. Protein comparisons were accomplished using MUSCLE[180].

Swarming study

Six strains were constructed using two different plasmids (pBBR1MCS and pYbiA) to determine what effect phage-encoded *ybiA* has on swarming (Table 4-1). Experimental data used from three biological and mechanical triplicate experiments using overnight cultures grown at 30℃ in 5 ml lysogeny broth (LB) and supplemented with 35 µg/ml chloramphenicol (Sigma Aldrich). Fresh M8 agar plates were poured and allowed to set for 60 min, then inoculated with 5 µl of ON in the center of the plate. Plates were stacked two high and incubated at 30℃ for 24 h followed by room temperature incubation for another 24 h. Plates were photographed at 24 and 48 h without automatic focus to ensure the scale did not change between plates. Images were analyzed using ImageJ software to measure the total area of the swarming bacteria on the plate. Determination of the mean and standard deviation for each sample was accomplished using Microsoft Excel.

FolA functionality

Four strains were constructed using *Escherichia coli* DH5α, *S. maltophilia* D1585 and two plasmids pBBR1MCS and pFolA to study the functionality of the DLP4 encoded *folA* (Table 4-1). Triplicate minimal inhibitory concentration (MIC) experiments used established protocols[243]. Overnight cultures were grown at 30℃ in 5 ml lysogeny broth (LB) with 35 µg/ml chloramphenicol. A 1:100 subculture was grown at 30℃ to an OD600 of 0.1 in Mueller-Hinton broth (MH) (approximately 2.5 hours) and used in 96 well plates containing a trimethoprim (TP) dilution series (MP Biomedicals). Following a 16-hour incubation, OD600 was observed using a Wallac 1420 VICTOR2 multilabel counter (PerkinElmer, Waltham, MA) and values were averaged using Excel. Statistical analysis was conducted using GraphPad Prism 7 (GraphPad Software Inc., San Diego, CA) to perform a two-way analysis of variance (ANOVA) with Sidak's multiple comparisons and P-values < 0.05 were documented.

RNA isolation and cDNA synthesis

A modified Epicentre Technologies: MasterPureTM RNA Purification protocol was used to isolate total RNA from *S. maltophilia*. Triplicate 5 ml cultures of *S. maltophilia* D1585 and lysogen D1585::DLP4 were grown in LB at 30˚C overnight and used for a 1:100 subculture into 10 ml LB at 30˚C for 4 hours (~3.0x108 CFU/ml). At the point of harvest, a 1.25 ml aliquot of ice-cold ethanol/phenol stop solution (5 % water-saturated phenol, pH <7) was added to the 10 ml culture. Cells were then pelleted by centrifugation at 5,000 g for 10 min and resuspended in 75 µl LB. A 25 µl aliquot of the suspension was transferred into three nuclease-free microcentrifuge tubes. A master mix was made using 3.5 µl Proteinase K (50 µg/µl) into 1 ml of Tissue and Cell Lysis Solution (Epicentre Technologies). A 300 µl aliquot of the master mix was added to each of the three nuclease-free tubes containing the resuspended bacterial culture and thoroughly mixed. The samples were incubated at 65˚C for 15 min with vortexing every 5 min. Following the 65˚C incubation, the cells were iced for 5 min, then 175 µl of MPC Protein Precipitation Reagent (Epicentre Technologies) was added to each tube and vigorously vortexed for 10 s. Particulates were pelleted by centrifugation for 10 min at >10,000 x g. An additional 25 µl of the MPC solution was added to the tubes which had a clear, small or loose pellet. Following centrifugation, the supernatant was transferred to a new nuclease-free tube with 500 µl isopropanol and inverted 30-40 times. The RNA was pelleted at 4˚C for 10 min at max rpm, followed by removal of the isopropanol layer. The pellet was then rinsed with 1 ml of 75 % EtOH, centrifuged briefly and resulting EtOH/isopropanol was removed with a pipette. One RNA pellet was resuspended in 100 µl nuclease-free water then transferred to the second and third tube to resuspend all three pellets in the 100 µl water. A 10 µl aliquot of 10x DNase I buffer (Ambion) and 10 units of RNase-free DNase I was added to the resuspended RNA solution and incubated at 37˚C for 10 min. The reaction was stopped with 5 µl of 50 mM EDTA, and 1 µl of SUPERase-IN (Ambion) was added. The resulting purified RNA was quantified then aliquoted into single-use tubes for storage at -80˚C.

Complementary DNA (cDNA) was synthesized using a modified protocol from GeneChip™ Expression Analysis Technical Manual with specific protocols for using the GeneChip™ (Thermofisher Scientific). The RNA concentrations were standardized to 500 ng/µl, and 3.5 µg of total RNA was used for the reactions. A 4 µl aliquot of random hexamers (Invitrogen) and 1 µl of dNTPs (10 mM) was dispensed into the tubes containing RNA, and the

final volume was adjusted to 12 µl with RNase-free water. This mix was incubated at 70˚C for 10 min, followed by 25˚C for 10 min and then chilled to 4˚C and briefly centrifuged. To this reaction mixture, 4 µl of 5x first strand buffer, 2 µl 0.1 M DTT, 1 µl SUPERase-IN and 1 µl SuperScript II (SSII). The solution was gently mixed and centrifuged, followed by these incubation steps: 25˚C 10 min, 37˚C 60 min, 42˚C for 60 min, inactivation of SSII at 70˚C for 10 min then chill to 4˚C. The resulting mixture was cleaned up with a QIAquick PCR Purification Kit (Qiagen) with a 40 µl elution.

Reverse transcription PCR

PCR analysis was conducted on each purified cDNA isolate using TopTaq DNA polymerase, buffer, and Q-solution (Qiagen), as well as primers specific to each gene of interest (Integrated DNA Technologies). Positive control primers were designed off of *S. maltophilia* D1585 *rpoD* (RNA polymerase sigma factor RpoD) and *proC* (proline biosynthetic gene). Gene-specific primers were designed from the DLP4 to detect *folA* (dihydrofolate reductase) and *ybiA* (N-glycosidase). The amount of cDNA used in each reaction was standardized to 200 ng. The resulting products were separated on a 1 % (wt/vol) agarose gel in 1x TAE (pH 8.0) and stained with ethidium bromide for visualization with a ChemiDoc MP imaging system and the Image Lab software (Bio-Rad).

RESULTS AND DISCUSSION

Isolation, host range, and morphology

Bacteriophage DLP4 (vB_SmaS-DLP_4) was isolated from asparagus soil in Edmonton, Alberta, Canada using clinical *Stenotrophomonas maltophilia* strain D1585. Electron microscopy of DLP4 reveals a long, noncontractile tail averaging 139 nm and capsid width and length of 63 and 92 nm respectively (Figure 4-1). The capsid width to length ratio is 1.46, classifying it as a B2 morphotype[223] of the *Siphoviridae* family and the *Caudovirales* order.

Host range analysis shows DLP4 is capable of lytic growth on 14 of 27 clinical *S. maltophilia* isolates (Table 4-2). Although DLP4 is most closely related to *Pseudomonas aeruginosa* phages AAT-1[244], PaMx2836 and PaMx7436 at the nucleotide level, host range analysis of DLP4 on *P. aeruginosa* strains showed it is not capable of infecting the *P. aeruginosa*

strains tested (Table 4-2). Plaque development by DLP4 occurs readily at 30˚C within 16 hours, forming diffuse plaques with irregular borders and a mean size of 0.8 ± 0.3 mm.

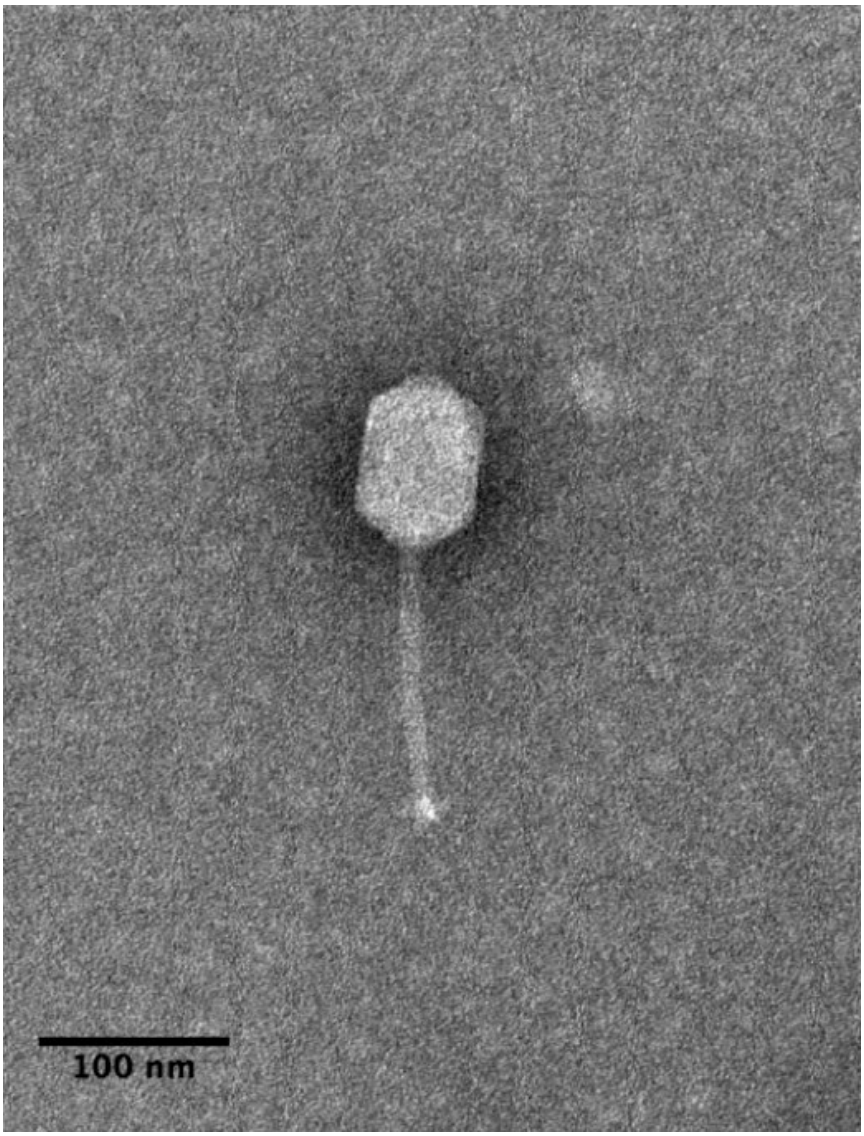

Figure 4-1: The DLP4 *Siphoviridae* phage of a B2 morphotype stained with 4 % uranyl acetate on a copper grid and viewed with a Philips/FEI transmission electron microscope. Scale bar represents 100 nm. Measurements of ten DLP4 phages give capsid width and length averages of 63 and 92 nm respectively, and tail length averages of 139 nm.

Table 4-2: Efficiency of plating host range of DLP4 against *S. maltophilia* and *P. aeruginosa* strains.

***S. maltophilia* strains**	**Efficiency of plating**	***P. aeruginosa* strains**	**Efficiency of plating**
101[c]	++	PA01	–
102[c]	++	HER1004	–
103[c]	+++	HER1012	–
152[c]	-	14715	–
155[c]	++++	Utah3	–
174[c]	-	Utah4	–
176[c]	-	14655	–
213[c]	-	6106	–
214[c]	-	pSHU-OTE	–
217[c]	-	D1606D[a,b]	–
218[c]	-	D1615C[a,b]	–
219[c]	++	D1619M[a,b]	–
230[c]	+	D1620E[a,b]	–
236[c]	-	D1623C[a,b]	–
242[c]	-	ENV003[a]	–
249[c]	-	ENV009[a]	–
278[c]	+	FC0507[a]	–
280[c]	++	R285	–
282[c]	++++	14672	–
287[c]	+		
446[c]	-		
667[c]	+		
D1585[a,b]	+++		
D1571[a,b]	-		
D1614[a,b]	-		
D1576[a,b]	++++		
D1568[a,b]	+++		

[a] Obtained from the Canadian *Burkholderia cepacia* complex Research Referral Repository.
[b] Cystic fibrosis patient isolate.
[c] Isolates from the Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services.
–, No sensitivity to phage; +, plaques at 10^{-2}; ++, clearing at 10^{-2}; +++, plaques at 10^{-4}; ++++, plaques at 10^{-6}.

Genome characterization

Restriction fragment length polymorphism (RFLP) analysis on purified gDNA was unsuccessful because 36 restriction enzymes tested failed to digest the genomic DNA. Restriction enzyme resistant DNA was also found with closely related *P. aeruginosa* phages PaMx2836 and PaMx7436. Although the panel of restriction enzymes was smaller (NdeI, HindIII, and EcoRI), the authors did notice a general trend of restriction enzyme resistant DNA in the other 17 phage B2 morphotypes studied in the paper[245]. DLP4 assembled into a 63,945 bp circular contig with a read coverage of 1928 and a 100 % Q40. The contig was confirmed with PCR using 15 primer sets followed by Sanger sequencing. The DLP4 genome has a GC content of 65 % and is predicted to encode 82 ORFs (Figure 4-2, Table 4-3). The modular arrangement of genes based on function shows distinct regions encoding genes involved in DNA replication and repair (blue), lysis (red), virion morphogenesis (mustard) and the YbiA operon (green) (Figure 4-2). Although DLP4 is confirmed to be a temperate phage capable of establishing a lysogenic infection within *S. maltophilia* D1585, the repressor and integrase could not be identified using BLASTp, CD-Search or Pfam. The genome sequence of DLP4 deposited in GenBank has the accession number MG018224.

DNA replication and repair module

DLP4 encodes 45 genes encoding proteins involved in DNA replication, repair and the generation and processing of nucleotides (BIT20_016 - 060). Within the module, some gene products which could be assigned a function include helicase (BIT20_031), DNA polymerase (BIT20_032), DNA binding protein (BIT20_039), Cas4 nuclease (BIT20_041), primase (BIT20_045), small terminase (BIT20_046), deoxynucleoside monophosphate kinase (BIT20_049), RNA pseudouridine synthase (BIT20_051), large terminase (BIT20_052), NrdA (BIT20_055), NrdB (BIT20_056), polynucleotide kinase (BIT20_059), DNA ligase (BIT20_060) and a protein with a conserved HIRAN domain (BIT20_019) (Figure 4-2, Table 4-3). Proteins with HIRAN domains have been shown to identify DNA damage and stalled replication forks[246], though the functionality of the DLP4 protein is currently unknown.

Identification of a Cas4 nuclease conserved domain within a DLP4-encoded protein warranted further investigation. Phage-encoded Cas4 nuclease homologs were previously identified in *Campylobacter jejuni* bacteriophages and were shown to be capable of inserting

new spacers into the CRISPR array of their host bacterium[247]. The spacers incorporated into the array were host-derived, suggesting the phages are making decoy spacers using the Cas4 nuclease to prevent the degradation of phage DNA[247]. BLASTp analysis of the Cas4 nuclease (BIT20_041) reveals this protein is also highly conserved within bacteriophages that infect a range of hosts such as *Acinetobacter, Xanthomonas, Pseudomonas,* and *Achromobacter*. To determine if DLP4 is using the putative Cas4 nuclease to incorporate new host-specific spacers like the *C. jejuni* phages, expression of *cas4* in the lysogen was examined using RT-PCR. No expression of *cas4* was observed during the lysogenic phase (data not shown). Attempts to identify CRISPR arrays within an *S. maltophilia* D1585 scaffold using CRISPRCasFinder (https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index) were unsuccessful.

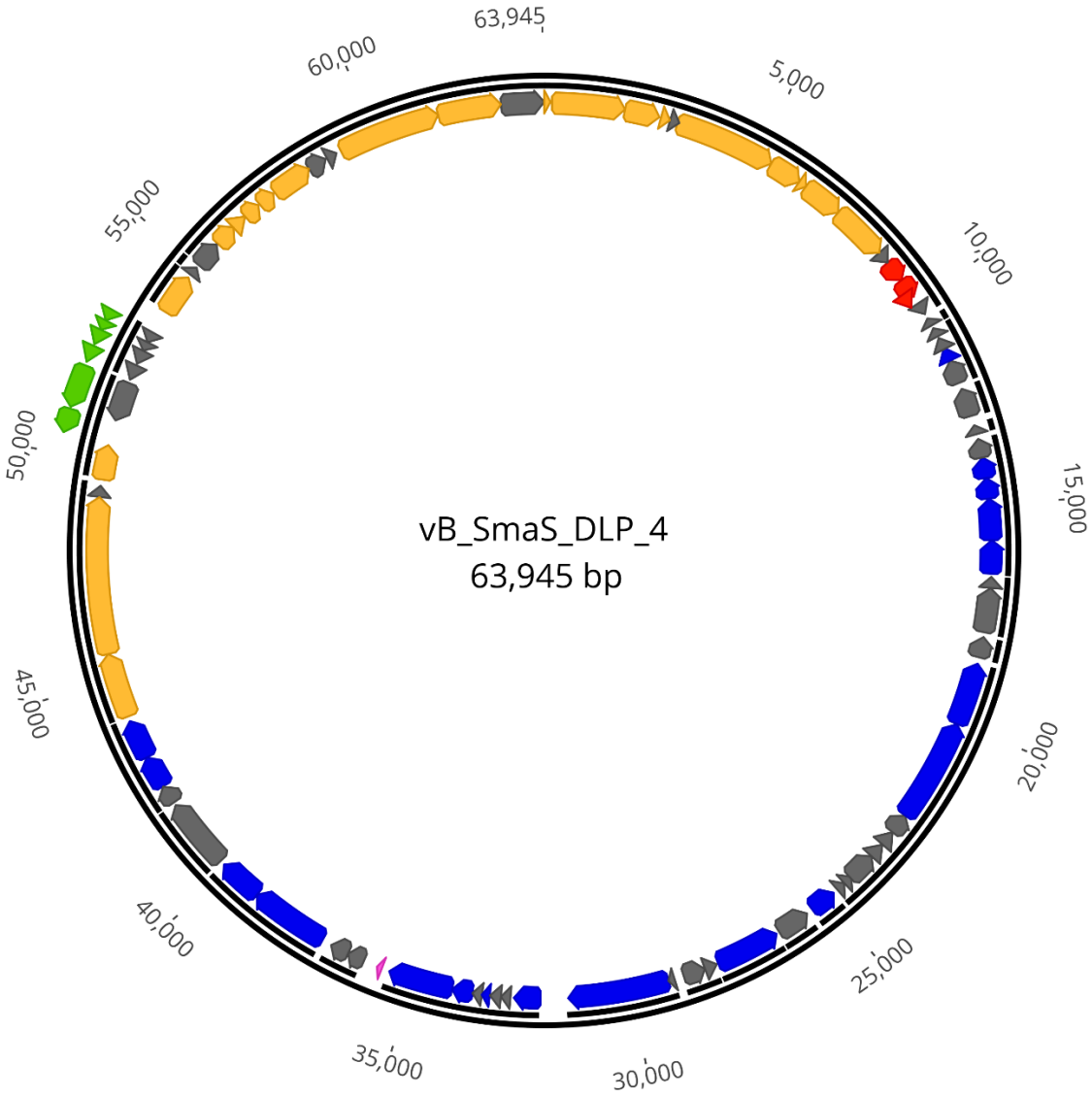

Figure 4-2: Circularized genome map of DLP4. The scale in bp is shown on outer periphery. Assigned functions for each predicted open reading frame are as follows: lysis; red, virion morphogenesis; mustard, DNA replication/repair; blue, hypothetical; grey, and YbiA operon; light green.

A cluster of genes encoded within the virion module encode proteins involved in the generation and processing of deoxyribonucleosides for their immediate use in phage DNA synthesis during the lytic cycle. DLP4 contains the genes *nrdA* and *nrdB* (BIT20_055/056 respectively) that encode the α2 and β2 subunits of ribonucleoside diphosphate reductase (RDR). The RDR heteromeric protein reduces ribonucleosides to deoxyribonucleosides, the first step in

the generation of deoxyribonucleoside triphosphates[248]. The next processing step of the resulting deoxyribonucleoside monophosphates (dNMP) is the addition of a phosphate group to make deoxyribonucleoside diphosphates (dNDP) using either ATP or dATP as the phosphate donor[249]. This step is carried out by the substrate-specific enzyme deoxyribonucleoside monophosphate kinase (BIT20_049) which is substrate specific to dGMP, dTMP, and 5-hydroxymethyl-dCMP[250].

Before all of the dCMP is phosphorylated to its dCDP form, some can be processed by the enzyme dCMP deaminase (deoxycytidylate deaminase: BIT20_025) to produce deoxyuridine monophosphate (dUMP). The dUMP product is the nucleotide substrate for thymidylate synthase (BIT20_026) which produces deoxythymidine monophosphate (dTMP). The thymidylate synthase reaction drives the concomitant conversion of 5,10-methylenetetrahydrofolate to dihydrofolate[251]. The dihydrofolate can then be reduced by dihydrofolate reductase (BIT20_024) into tetrahydrofolic acid, which is processed by serine transhydroxymethylase to regenerate the 5,10-methylenetetrahydrofolate. This enzyme was not identified in DLP4, but the enzyme serine:glyoxylate aminotransferase (BIT20_002) is encoded which catalyzes the reversible reaction of glycine and hydroxypyruvate to produce glyoxylate and L-serine (KEGG reaction: R00588). This L-serine could then be used by the hosts' serine transhydroxymethylase (encoded by *glyA* in *S. maltophilia*) to regenerate the 5,10-methylenetetrahydrofolate for the dUMP to dTMP reaction catalyzed by thymidylate synthase.

Table 4-3: Annotations of DLP4 genome using BLASTp analysis of translated coding regions.

Gene	Coding region	Length (AA)	Strand	Start codon	Putative function	BLASTp result	E value	Identity	Accession
1	5 - 1135	377	+	ATG	hypothetical protein	virion structural protein [*Pseudomonas* phage AAT-1]	3.00E-13	62%	AME18051.1
2	1117 - 2799	561	+	ATG	serine--glyoxylate aminotransferase	putative virion structural protein [*Pseudomonas* phage PaMx74]	0	63%	YP_009199471.1
3	2799 - 3611	271	+	ATG	FAD/FMN dehydrogenase	putative virion structural protein [*Pseudomonas* phage PaMx74]	7.00E-151	73%	YP_009199472.1
4	3624 - 3857	78	+	ATG	virion structural protein	putative virion structural protein [*Pseudomonas* phage PaMx74]	1.00E-39	83%	YP_009199473.1
5	3854 - 4057	68	+	GTG	hypothetical protein	hypothetical protein PaMx74_35 [*Pseudomonas* phage PaMx74]	7.00E-23	64%	YP_009199474.1
6	4044 - 6359	772	+	ATG	virion structural protein	putative virion structural protein [*Pseudomonas* phage PaMx74]	0	58%	YP_009199475.1
7	6359 - 7138	260	+	ATG	tail assembly protein	tail assembly protein [*Xylella* phage Salvo]	2.00E-82	51%	AHB12242.1
8	7142 - 7306	55	+	ATG	tail assembly protein	tail assembly protein [*Xylella* phage Sano]	1.00E-04	37%	AHB12066.1
9	7316 - 8260	315	+	ATG	tail assembly protein	tail assembly protein [*Xylella* phage Salvo]	1.00E-90	51%	AHB12240.1
10	8263 - 9546	428	+	GTG	hypothetical protein	tail fiber protein [*Xylella* phage Salvo]	7.00E-12	41%	AHB12239.1
11	9543 - 9848	102	+	GTG	hypothetical protein	hypothetical protein AAT1_02032 [*Pseudomonas* phage AAT-1]	7.00E-39	64%	AME18058.1

12	9845 - 10339	165	+	ATG	endolysin	putative endolysin [*Pseudomonas* phage PaMx74]	3.00E-92	83%	YP_009199477.1
13	10350 - 10826	159	+	ATG	i-spanin	virion structural protein [*Pseudomonas* phage PaMx28]	8.00E-59	63%	YP_009210650.1
14	10639 - 11013	125	+	GTG	o-spanin	putative o-spanin [*Pseudomonas* phage AAT-1]	2.00E-51	68%	ANN44564.1
15	11010 - 11285	92	+	GTG	hypothetical membrane protein	putative membrane protein [*Pseudomonas* phage PaMx74]	2.00E-37	67%	YP_009199480.1
16	11417 - 11593	59	-	ATG	hypothetical protein	hypothetical protein PaMx74_42 [*Pseudomonas* phage PaMx74]	4.00E-04	40%	YP_009199481.1
17	11661 - 11891	77	-	ATG	hypothetical protein	hypothetical protein PaMx74_43 [*Pseudomonas* phage PaMx74]	3.00E-14	63%	YP_009199482.1
18	11918 - 12202	95	-	ATG	hypothetical protein	hypothetical protein AAT1_02038 [*Pseudomonas* phage AAT-1]	4.00E-04	37%	AME18064.1
19	12186 - 12491	102	-	ATG	HIRAN domain protein	putative HIRAN domain-containing protein [*Pseudomonas* phage PaMx28]	2.00E-27	57%	YP_009210657.1
20	12491 - 13048	186	-	ATG	hypothetical protein	hypothetical protein PaMx74_46 [*Pseudomonas* phage PaMx74]	5.00E-58	60%	YP_009199485.1
21	13171 - 13839	223	-	ATG	hypothetical protein	hypothetical protein PaMx74_47 [*Pseudomonas* phage PaMx74]	7.00E-62	50%	YP_009199486.1
22	14040 - 14264	75	-	ATG	hypothetical protein	hypothetical protein PaMx74_48 [*Pseudomonas* phage PaMx74]	7.00E-25	76%	YP_009199487.1
23	14386 - 14832	149	-	GTG	hypothetical protein	hypothetical protein PaMx74_49 [*Pseudomonas* phage PaMx74]	3.00E-70	75%	YP_009199488.1
24	14817 - 15305	163	-	ATG	FolA/DHFR	putative dihydrofolate reductase [*Pseudomonas* phage PaMx74]	3.00E-59	61%	YP_009199489.1

25	15290 - 15766	159	-	ATG	deoxycytidylate deaminase	putative dCMP deaminase [*Pseudomonas* phage PaMx74]	6.00E-70	70%	YP_009199490.1
26	15766 - 16707	314	-	GTG	thymidylate synthase	thymidylate synthase [*Pseudomonas* phage AAT-1]	1.00E-161	71%	AME18072.1
27	16704 - 17477	258	-	ATG	nucleotide pyrophosphohydrolase	putative nucleotide pyrophosphohydrolase [*Pseudomonas* phage AAT-1]	3.00E-106	63%	AME18073.1
28	17553 - 17783	77	-	ATG	hypothetical protein				
29	17786 - 18805	340	-	ATG	hypothetical protein	hypothetical protein PaMx28_55 [*Pseudomonas* phage PaMx28]	0	74%	YP_009210667.1
30	18920 - 19384	155	-	ATG	hypothetical protein	hypothetical protein PaMx28_56 [*Pseudomonas* phage PaMx28]	2.00E-42	78%	YP_009210668.1
31	19526 - 21004	493	-	GTG	helicase	putative helicase [*Pseudomonas* phage PaMx74]	0	80%	YP_009199496.1
32	21001 - 23376	792	-	ATG	DNA polymerase	putative DNA polymerase [*Pseudomonas* phage PaMx74]	0	77%	YP_009199498.1
33	23405 - 23851	149	-	GTG	hypothetical protein	hypothetical protein AAT1_02054 [*Pseudomonas* phage AAT-1]	6.00E-60	57%	AME18080.1
34	23848 - 24237	130	-	TTG	hypothetical protein	hypothetical protein PaMx74_61 [*Pseudomonas* phage PaMx74]	3.00E-60	75%	YP_009199500.1
35	24234 - 24578	115	-	ATG	hypothetical protein	hypothetical protein PaMx74_63 [*Pseudomonas* phage PaMx74]	2.00E-54	75%	YP_009199502.1
36	24575 - 25258	228	-	ATG	hypothetical protein	hypothetical protein PaMx74_65 [*Pseudomonas* phage PaMx74]	1.00E-89	56%	YP_009199504.1
37	25248 - 25448	67	-	GTG	hypothetical protein	hypothetical protein PaMx74_66 [*Pseudomonas* phage PaMx74]	8.00E-19	58%	YP_009199505.1

38	25445 - 25672	76	-	ATG	hypothetical protein				
39	25793 - 26404	204	-	ATG	DNA binding protein	putative DNA binding protein [*Pseudomonas* phage PaMx28]	2.00E-108	78%	YP_009210678.1
40	26538 - 27317	260	-	ATG	hypothetical protein	hypothetical protein PaMx28_69 [*Pseudomonas* phage PaMx28]	1.00E-76	59%	YP_009210681.1
41	27390 - 28868	493	-	ATG	Cas4 nuclease	PD-(D/E)XK nuclease superfamily protein	0	72%	YP_009210682.1
42	28918 - 29181	88	-	ATG	hypothetical protein	hypothetical protein PaMx74_72 [*Pseudomonas* phage PaMx74]	5.00E-43	81%	YP_009199511.1
43	29178 - 29702	175	-	ATG	hypothetical protein	hypothetical protein PaMx28_72 [*Pseudomonas* phage PaMx28]	6.00E-23	43%	YP_009210684.1
44	29904 - 30074	57	+	ATG	hypothetical protein	hypothetical protein PaMx74_74 [*Pseudomonas* phage PaMx74]	6.00E-16	72%	YP_009199513.1
45	30071 - 32392	774	+	GTG	primase	putative primase/polymerase [*Pseudomonas* phage PaMx74]	0	85%	YP_009199514.1
46	33021 - 33638	206	+	GTG	small terminase	terminase large subunit [*Pseudomonas* phage AAT-1]	9.00E-108	77%	AME18098.1
47	33701 - 33919	73	+	ATG	hypothetical protein				
48	33929 - 34171	81	+	ATG	hypothetical protein	hypothetical protein PaMx74_02 [*Pseudomonas* phage PaMx74]	1.00E-27	67%	YP_009199441.1
49	34168 - 34362	65	+	ATG	deoxynucleoside monophosphate kinase	hypothetical protein [*Lysobacter* sp. Root667]	1.00E-08	56%	WP_056102216.1
50	34359 - 34577	73	+	ATG	hypothetical protein	hypothetical protein [*Lysobacter capsici*]	2.00E-06	47%	WP_057921118.1

51	34581 - 35045	155	+	GTG	RNA Pseudouridine synthase				
52	35045 - 36514	490	+	ATG	large terminase	putative terminase large subunit [*Pseudomonas* phage PaMx74]	0	80%	YP_009199443.1
53	37156 - 37560	135	+	ATG	hypothetical protein	hypothetical protein PaMx74_06 [*Pseudomonas* phage PaMx74]	6.00E-07	34%	YP_009199445.1
54	37557 - 37991	145	+	ATG	hypothetical protein				1
55	38200 - 40020	607	+	ATG	nrdA	nrdA [uncultured Mediterranean phage uvMED]	0	51%	BAQ94146.1
56	40028 - 41020	547	+	ATG	nrdB	nrdB [uncultured Mediterranean phage uvMED]	5.00E-126	54%	BAR25383.1
57	41131 - 42771	331	+	ATG	hypothetical protein	hypothetical protein PaMx74_08 [*Pseudomonas* phage PaMx74]	4.00E-169	51%	YP_009199447.1
58	42840 - 43274	145	+	ATG	hypothetical protein	hypothetical protein PaMx74_10 [*Pseudomonas* phage PaMx74]	1.00E-06	40%	YP_009199449.1
59	43271 - 44059	263	+	GTG	polynucleotide kinase	hypothetical protein [*Mitsuaria chitosanitabida*]	3.00E-38	46%	WP_067070380.1
60	44059 - 44955	299	+	ATG	DNA ligase	putative DNA ligase [*Pseudomonas* phage PaMx74]	2.00E-106	57%	YP_009199450.1
61	45042 - 46562	507	+	ATG	portal protein	structural protein [*Pseudomonas* phage AAT-1]	0	78%	AME18030.1
62	46562 - 50131	1190	+	ATG	minor head protein	morphogenesis protein [*Pseudomonas* phage PaMx28]	0	75%	YP_009210622.1
63	50133 - 50405	91	+	ATG	hypothetical protein	hypothetical protein PaMx28_11 [*Pseudomonas* phage PaMx28]	1.00E-42	80%	YP_009210623.1

64	50549 - 51334	262	+	ATG	virion structural protein	virion structural protein [*Pseudomonas* phage AAT-1]	8.00E-146	78%	AME18033 .1
65	51384 - 51869	162	-	GTG	YbiA	putative YbiA-like protein [*Pseudomonas* phage PaMx28]	2.00E-81	72%	YP_009210 625.1
66	51909 - 52817	303	-	ATG	hypothetical protein	hypothetical protein PaMx74_15 [*Pseudomonas* phage PaMx74]	3.00E-70	47%	YP_009199 454.1
67	52903 - 53280	126	-	ATG	hypothetical protein	hypothetical protein AAT1_02012 [*Pseudomonas* phage AAT-1]	2.00E-19	49%	AME18037 .1
68	53292 - 53621	110	-	ATG	hypothetical protein	hypothetical protein [Enterobacteriaceae]	0.034	35%	WP_04434 7588.1
69	53636 - 53848	71	-	ATG	hypothetical protein	hypothetical protein AAT1_02013 [*Pseudomonas* phage AAT-1]	9.00E-20	71%	AME18038 .1
70	53856 - 54095	80	-	ATG	hypothetical protein				
71	54657 - 55589	311	+	ATG	major head protein	major head protein [*Pseudomonas* phage PaMx28]	0	87%	YP_009210 631.1
72	55657 - 55890	78	+	ATG	hypothetical protein				
73	55957 - 56568	204	+	ATG	hypothetical protein	hypothetical protein PaMx74_20 [*Pseudomonas* phage PaMx74]	5.00E-47	49%	YP_009199 459.1
74	56591 - 57115	175	+	ATG	virion structural protein	putative virion structural protein [*Pseudomonas* phage PaMx74]	1.00E-84	72%	YP_009199 461.1
75	57117 - 57485	123	+	ATG	virion structural protein	putative virion structural protein [*Pseudomonas* phage PaMx74]	2.00E-46	61%	YP_009199 462.1
76	57487 - 57879	131	+	GTG	virion structural protein	virion structural protein [*Pseudomonas* phage PaMx28]	5.00E-72	80%	YP_009210 635.1

77	57892 - 58314	141	+	ATG	tail terminator protein	putative tail terminator protein [*Pseudomonas* phage PaMx74]	6.00E-92	93%	YP_009199464.1
78	58337 - 59278	314	+	ATG	major tail structural protein	major tail structural protein [*Pseudomonas* phage AAT-1]	0	80%	AME18046.1
79	59283 - 59729	149	+	ATG	hypothetical protein	hypothetical protein AAT1_02022 [*Pseudomonas* phage AAT-1]	3.00E-54	61%	AME18047.1
80	59768 - 60013	82	+	ATG	hypothetical protein	hypothetical protein PaMx74_28 [*Pseudomonas* phage PaMx74]	3.00E-46	90%	YP_009199467.1
81	60117 - 62474	786	+	ATG	tape measure protein	putative tail tape measure protein [*Pseudomonas* phage PaMx74]	0	74%	YP_009199468.1
82	62490 - 63944	485	+	ATG	tail fiber structural protein	tail fiber structural protein [*Pseudomonas* phage MP1412]	3.00E-98	39%	YP_006561077.1

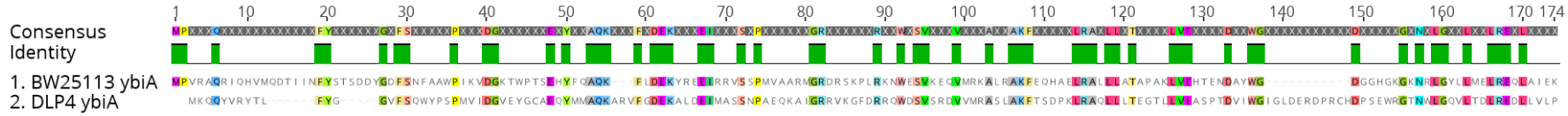

Figure 4-3: MUSCLE protein alignment of DLP4 YbiA against *Escherichia coli* BW25113 YbiA. Functional residues are located at positions 48 (E), 92 (W), 133 (D) and 136 (W), are present.

.

Lysis module

The lysis module (Figure 4-2, Table 4-3) of this bacteriophage directly follows virion morphogenesis genes and is composed of five ORFs (BIT20_011- 015). The first gene in the lysis module encodes the holin protein (BIT20_011) that is predicted to be a class II holin due to the presence of two transmembrane domains. The next gene encodes a predicted endolysin (BIT20_012) with a conserved D-alanyl-D-alanine carboxypeptidase identified by CD-Search. This domain is also featured in DacA of *E. coli* K-12, which is responsible for trimming the carboxy-terminal D-alanyl residues from the peptidoglycan pentapeptides in *E. coli* [252]. The next two ORFs (BIT20_013/014) following encode i- and o- spanins respectively. The two spanins associate together in the periplasm and physically link the inner membrane to the outer membrane. The i-spanin is tethered to the inner membrane near the N-terminal domain though a transmembrane region and the C-terminal domain associates with the o-spanin in the periplasm[253]. The o-spanin contains a modified N-terminal Cys residue which has added fatty acid and diacylglycerol groups to anchor the lipoprotein to the outer membrane, allowing the C-terminal domain to associate with the C-terminal domain of the i-spanin[253]. The final ORF of the lysis cassette (BIT20_015) is predicted to encode a protein with a transmembrane domain. A conserved domain belonging to the SpsE protein superfamily, and more specifically to the NeuB_NnaB (TIGR03569) family, was identified within the gp15 protein. The NeuB_NnaB family consists of functional N-acetylneuraminate synthase proteins which produce N-acetylneuraminic acid (NANA), a sialic acid that is used by bacterial pathogens to hide from their mammalian hosts[254]. It is unclear why this protein is encoded within the lysis module of the genome, but DLP4 is not the only bacteriophage to encode this protein in the lysis module. Three *Pseudomonas* phages (PaMx7436, AAT-135, PaMx2836) and one *Xanthomonas* phage (Xoo-sp246) encode the same set of proteins as DLP4.

Virion morphogenesis module

The virion morphogenesis module of DLP4 encodes 27 ORFs, and BLASTp results against the translated proteins provided close matches for all but one protein, BIT20_072 (Figure 4-2, Table 4-3). Of the 27 predicted proteins within the module, 11 have high protein sequence identity to the *Pseudomonas* phage PaMx74. These proteins include three hypotheticals, a tape measure protein (BIT20_081), putative tail terminator protein (BIT20_077), a putative

FAD/FMD – containing dehydrogenase (BIT20_003), four putative virion structural proteins (BIT20_004, BIT20_006, BIT20_074, and BIT20_075) and serine-glyoxylate aminotransferase (BIT20_002). The phage AAT-1 BLASTp results that are close to five DLP4 proteins include two hypotheticals, the portal protein (BIT20_061), a putative structural protein (BIT20_064) and the major tail structural protein (BIT20_078). The tail assembly proteins of DLP4 (BIT20_007 - 009) are most similar to two *Xylella* phages named Salvo and Sano[255]. The ORF following the tail assembly genes of DLP4 encodes a tail fiber protein (BIT20_010) which is similar to the *Xylella* phage Salvo, though CD-Search does not identify a tail fiber domain in this protein. BLASTp also identified similar proteins in phage PaMx28: the major and minor head proteins (BIT20_071 and BIT20_062 respectively), a hypothetical protein (BIT20_063) and one virion structural protein (BIT20_076). When looking at the virion morphogenesis module, it appears there may be an abundance of recombination between phages PaMx74, PaMx28, Salvo, Sano and AAT-1 to result in DLP4 encoding so many distinct regions of identity at the protein level to each of the five phages.

YbiA operon

Within the virion morphogenesis region, there is an insert of approximately 2,750 bp in the reverse frame encoding six genes (BIT20_065 –070) (Figure 4-2, Table 4-3). The insert also exhibits a drop in GC content from the surrounding 66 % down to 61 %. The operon is under the control of a single promoter located 65 bp upstream of the first gene of the insert, BIT20_070. The six genes in this operon are also found within three other phages (*Xanthomonas* phage Xoo-sp2[256] and *Pseudomonas* phages PaMx28 and AAT-1) in the same orientation. The *Pseudomonas* phage PaMx74 contains a single gene (PaMx74_15) which is similar to BIT20_066 found within the DLP4 operon, though the surrounding genes of PaMx74_15 do not have any identity to the rest of the genes within the DLP4 operon.

The BIT20_070 gene product did not show any similar proteins in the NCBI database with a BLASTp search, and I-TASSER analysis did not give significant results with high confidence. The gene product of BIT20_069 showed identity to the hypothetical protein AAT1_02013 from *Pseudomonas* phage AAT-1, but no conserved domains were discovered using CD-Search. BIT20_068 does not have any similar proteins identified with BLASTp when limited to viruses, but with no database restrictions, the BIT20_068 protein shows identity to a

hypothetical multiple-species protein within Enterobacteriaceae. BLASTp analysis on the BIT20_67 gene product shows protein identity to the hypothetical protein AAT1_02012 belonging to *Pseudomonas* phage AAT-1 (73% identity, 2.0E-19, 49% coverage). I-TASSER structural prediction of the protein found similar possible structures. The top result (TM-score 0.705 and coverage 0.952) provided by I-TASSER was an anti-sigma factor antagonist with the protein data bank ID of 3ZTA. This database entry is for a *Moorella thermoacetica* protein shown to be involved in the bacterial stressosome, which is responsible for controlling secondary messenger signaling[257]. The BIT20_066 gene product shows high identity to the hypothetical protein PaMx74_15 from *Pseudomonas* phage PaMx74. I-TASSER results for this protein predicts it is structurally similar to the protein data bank ID 1FOH (TM-score 0.875 and coverage 0.937), a phenol hydrolase from *Trichosporon cutaneum* within the Fungi kingdom.

The final gene in the operon (BIT20_065) is predicted to encode a putative YbiA-like protein from *Pseudomonas* phage PaMx28. YbiA is responsible for increasing the swarming phenotype of *E. coli* K-12[258]. I-TASSER analysis of the DLP4 YbiA-like protein produced a TM-score of 0.896 and coverage of 0.913 to the YbiA protein of *E. coli*. A MUSCLE protein alignment between the two proteins shows a breakdown in sequence identity at the protein level, though known functional domains of YbiA (residues 48, Glu; 92, Trp; 133, Asp and 136, Trp)[259] are still present in the DLP4 protein sequence. The numbering for the last three residues changes to 92, 133 and 136 within the alignment (Figure 4-3). Analysis of the annotation results suggests the operon encodes moron genes which may help the host cell respond to environmental stressors.

Investigation of swarming phenotype

Identification of the *ybiA* gene within a small operon of DLP4 raised questions about the ability of the translated YbiA-like protein to affect swarming within D1585. The ability of *S. maltophilia* to consistently swarm in lab conditions under debate, with conflicting reports presented in the current literature[27-29]. One study showed *S. maltophilia* capable of what looked to be swarming, though it was found to be flagella-independent translocation in the presence of extracellular fatty acids[28]. Researchers observed the translocated cells for the presence of flagella, required for the swarming phenotype, but they found flagella were absent[28]. Observation of swarming plates of wild-type D1585, the lysogen and D1585 containing DLP4 pYbiA and

empty vector control (pBBR1MCS) did not show a swarming phenotype (data not shown). The predicted structural folding of DLP4 YbiA to *E. coli* YbiA suggests the DLP4 *ybiA* might complement a *ybiA*- *E. coli* mutant; therefore, swarming experiments were conducted in *E. coli* strains BW25113 and *ybiA*770(del)::kan. The swarming results suggest the DLP4 encoded YbiA can complement the *ybiA*- knockout to wild type swarming levels (Figure 4-4). Repeated swarming assays could not reduce the considerable variation observed within biological and mechanical replicates, though it is important to note the variation was found in all strains studied. RT-PCR of the lysogen and wild type D1585 showed *ybiA* is expressed in the lysogen (Figure 4-5), but it does not have the same effect on the swarming phenotype of D1585 as it did with the *E. coli* strains.

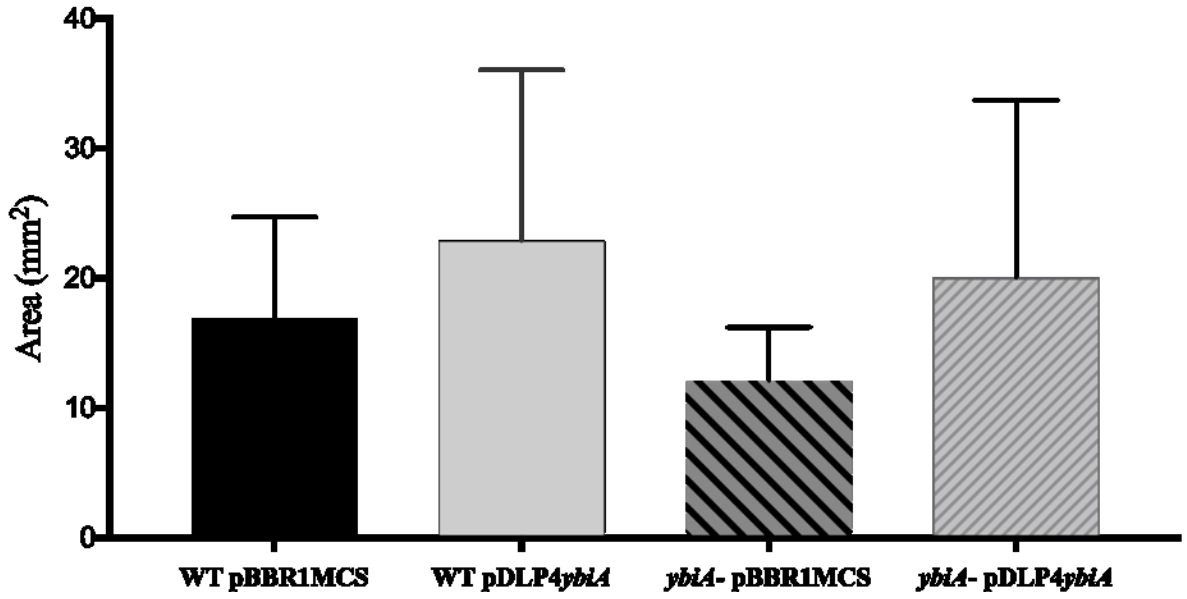

Figure 4-4: Complementation of *E. coli* BW25113 *ybiA*$^{-}$ with DLP4 *ybiA* restores swarming phenotype. Image data from three biological and mechanical triplicate swarming experiments was used to measure swarm area using ImageJ. Area averages and standard deviation were calculated and graphed in GraphPad Prism.

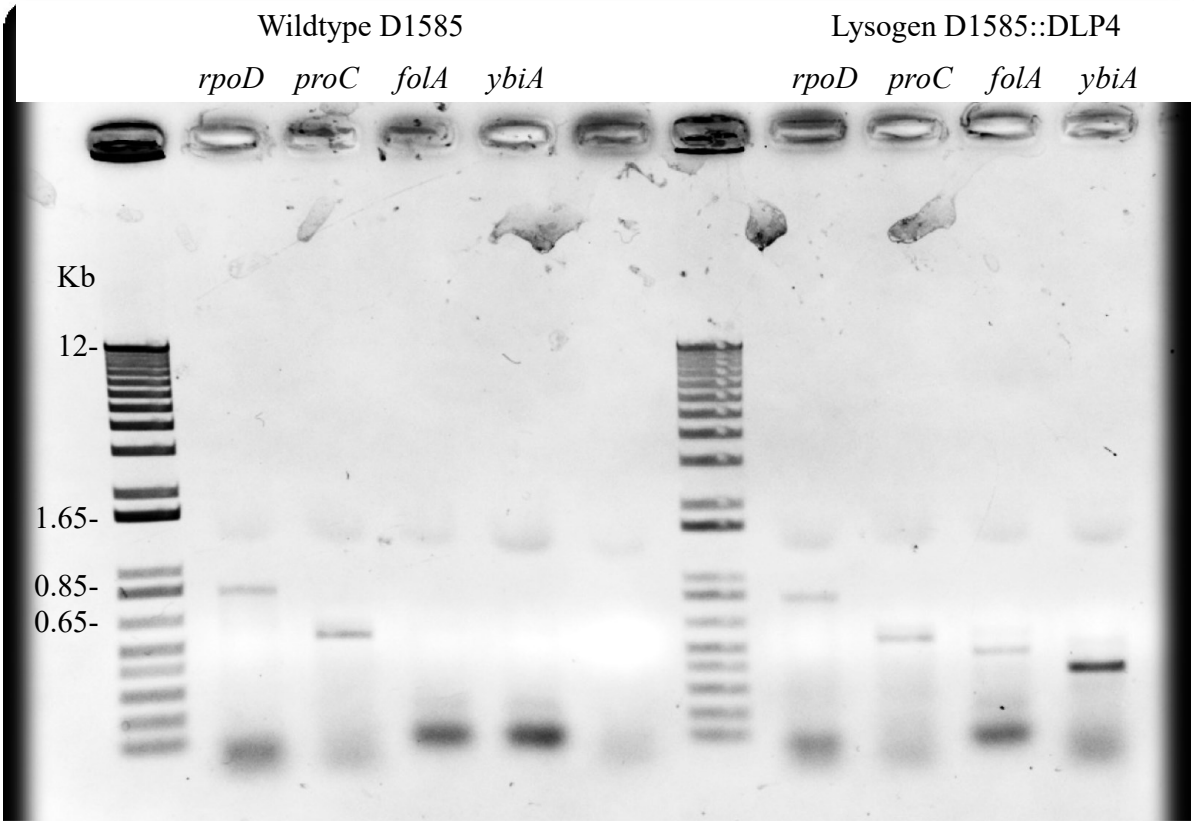

Figure 4-5: Reverse transcription PCR detects expression from the *folA* and *ybiA* genes in the D1585::DLP4 lysogen compared to wild type D1585 control during exponential phase. Positive controls are the exponential sigma factor-encoding gene *rpoD* (~ 850 bp) and Pyrroline-5-carboxylate reductase-encoding gene *proC* (~ 600 bp). Expression of *folA* (~ 500 bp) and *ybiA* (~ 400 bp) is observed in the lysogen when separated on a 1 % agarose gel. An Invitrogen 1 kb plus DNA ladder was used for size comparison.

Investigation of dihydrofolate reductase functionality

The discovery that DLP4 encodes *folA* (BIT20_024) was surprising at first, but in the context of the surrounding genes involved in deoxyribonucleoside generation, its place is fitting. I-TASSER analysis of DLP4 FolA predicted its structural similarity to *Bacillus anthracis* DHFR (FolA: TM-score 0.919, coverage 0.963). FolA is responsible for trimethoprim resistance in bacteria, so it was important to investigate if the DLP4 encoded *folA* produces a functional FolA causing lysogenic conversion of the host bacterium. Comparing the resistance profile of D1585 to D1585::DLP4, there is a statistically significant increase in TP resistance at 391 (*P*-value 0.0003), 586 (*P*-value <0.0001) and 781 (*P*-value 0.004) µg/ml concentrations (Figure 4-6). To confirm the DLP4 FolA is functional, the *folA* gene was cloned into the pBBR1MCS plasmid and expressed in *E. coli* DH5α and *S. maltophilia* D1585 against an empty vector control in

varying concentrations of trimethoprim (Figure 4-7). The results confirm the DLP4 FolA is functional, causing an increase in resistance to TP when expressed. The FolA protein increased the trimethoprim LD_{90} for DH5α-pfolA strain to 3,000 µg/ml from the <12 µg/ml observed with the empty vector control.

To confirm increased TP resistance observed for the lysogen was explicitly due to the expression of *folA* in the lysogen, reverse transcription PCR (RT-PCR) was used. Positive controls for the RT-PCR were gene-specific primers designed against D1585 housekeeping genes *rpoD* (σ70) and *proC* (pyrroline-5-carboxylate reductase)[260]. The RT-PCR results show *folA* is expressed during the lysogenic cycle (Figure 4-5), which supports the observation of increased TP resistance of the lysogen compared to wild type control.

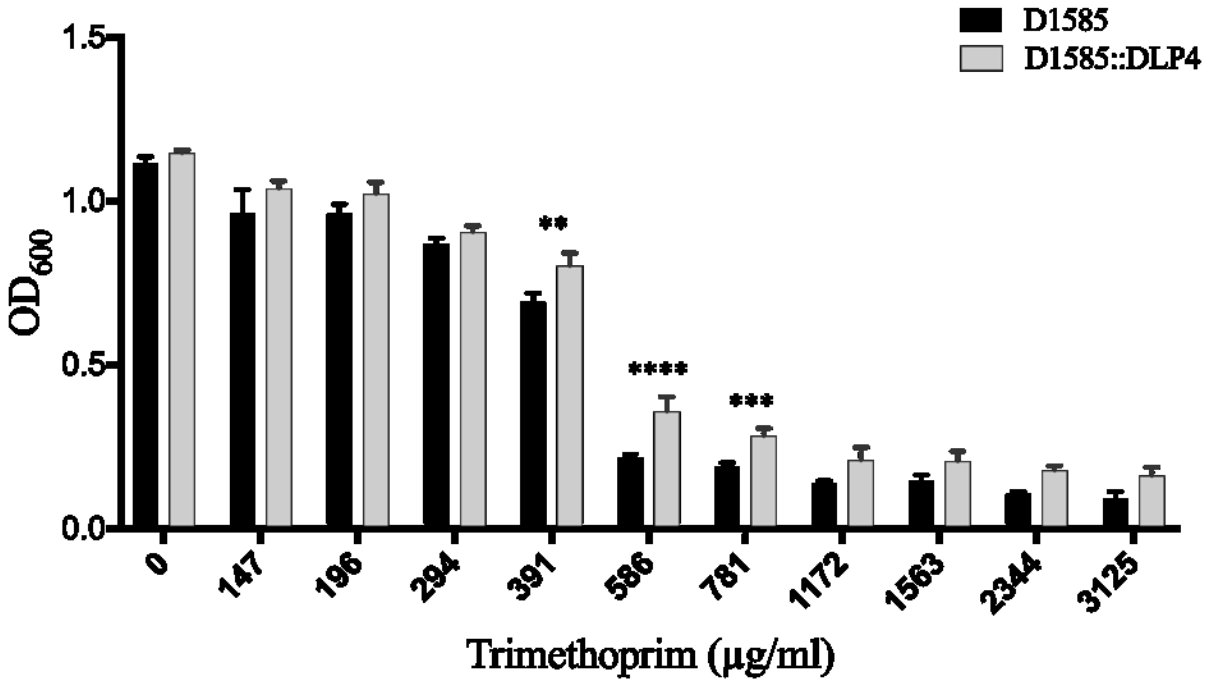

Figure 4-6: Trimethoprim resistance of D1585::DLP4 lysogen increases compared to wild type control. Assay was completed in biological and mechanical triplicate. Two-way ANOVA with Sidak's multiple comparisons test was performed on the MIC data, and statistical significance is represented as: ****, P < 0.0001; ***, P <0.001, ** P <0.01.

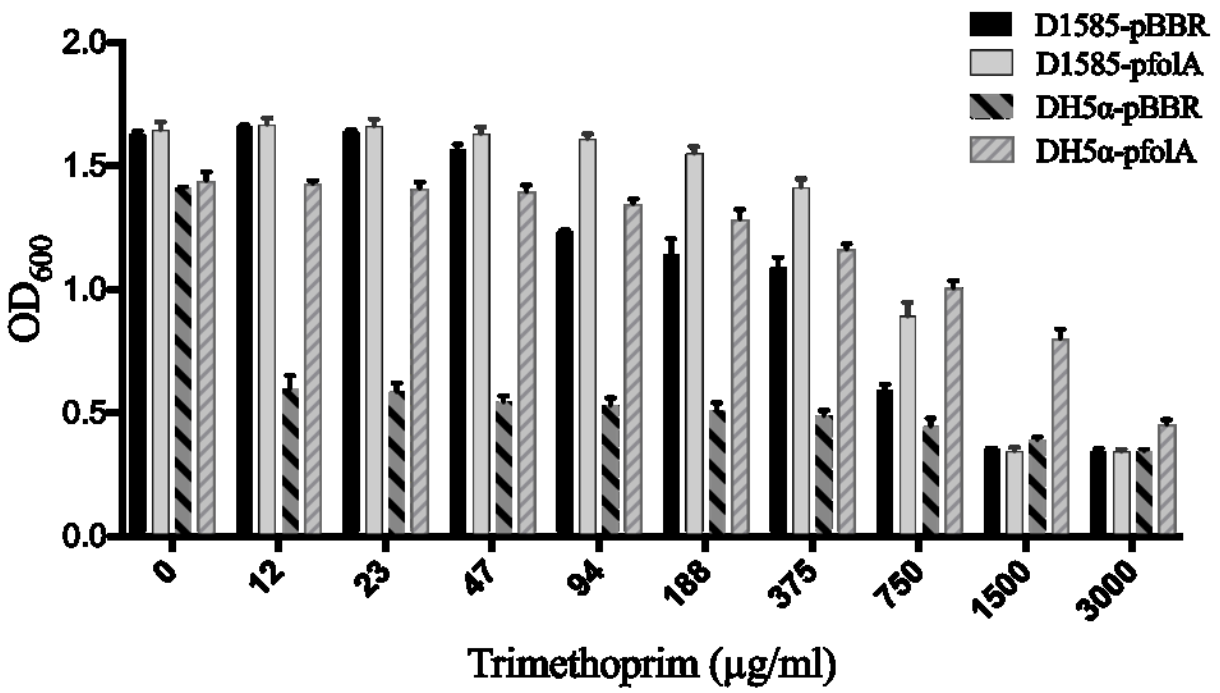

Figure 4-7: Increased trimethoprim resistance is due to DLP4 *folA* expression. Trimethoprim resistance in *E. coli* DH5α increases from <12 μg/ml to an LD_{90} of 3,000 μg/ml when DLP4 *folA* is expressed from the pBBR1MCS backbone compared to empty vector control

CONCLUSIONS

Characterization of the novel temperate phage DLP4 shows it features a restriction enzyme resistant genome 63,945 bp in size encoding 82 potential ORFs. The GC content of the DLP4 genome is found to be reflective of the host GC content of 65 %. DLP4 encodes a near complete deoxynucleoside conversion and salvage pathway including a functional dihydrofolate reductase which was shown to be functional and expressed in the lysogen. The DLP4 encoded YbiA operon has a functional YbiA protein that is required for the swarming phenotype of *E. coli* and is expressed during the lysogenic cycle, though no swarming was noticed for *S. maltophilia* D1585. The operon also encodes proteins that may be involved in bacterial stress response, such as a putative phenol hydrolase and an anti-sigma factor antagonist involved in the bacterial stressosome. A putative Cas4 nuclease is encoded within the DNA replication and repair module of DLP4, though the role of this protein in the DLP4 infection cycle is unknown, and it is not expressed during the lysogenic cycle. Although DLP4 is not suitable for therapeutic use due to the presence of these moron genes, this is the fifth example of a *Stenotrophomonas*

maltophilia phage carrying problematic moron genes which could lead to the lysogenic conversion of susceptible SMC members[154,156,159]. More research into temperate *S. maltophilia* bacteriophages could help elucidate the potential role these phages play in the virulence and antibiotic resistance of *S. maltophilia* isolates.

ACKNOWLEDGEMENTS

The authors would like to thank David Speert of the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR) for the kind gift of numerous *S. maltophilia* strains. The authors would also like to thank the Provincial Laboratory for Public Health— North (Microbiology), Alberta Health Services, for access to *S. maltophilia* clinical isolates.

Chapter 5

Temperate siphovirus DLP5 is the type species of genus *Delepquintavirus*

Portions of this chapter have been submitted as:

Peters, D. L. and J. J. Dennis. 2018. Complete genome sequence of temperate *Stenotrophomonas maltophilia* bacteriophage DLP5. Genome Announcements. **6** (9) e00073-18.

OBJECTIVES

The objectives of this research were to characterize the DLP5 genome, identify its lifestyle, and determine if lysogenic conversion is occurring in the lysogen.

MATERIALS AND METHODS

Bacterial strains and growth conditions

A total of 27 clinical *S. maltophilia* strains were used for host range analysis of DLP5. The Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR; Vancouver, BC) donated the strains D1585, D1571, D1614, D1576, and D1568 which were used for phage hunting. The Provincial Laboratory for Public Health – North (Microbiology), Alberta Health Services donated an additional 22 strains to expand the host range panel. Most strains were grown aerobically for 18 h at 30 ˚C on Luria-Bertani solid media or in LB broth with shaking at 225 RPM. The strain D1614 required a longer incubation on solid media (36 h) to obtain visible single colonies.

Phage isolation, propagation, host range analysis, and electron microscopy

DLP5 (vB_SmaS_DLP_5) was isolated from garden soil from Edmonton, Alberta, Canada using strain D1614 and a previously described extraction protocol[163]. Once isolated, the phage was propagated with soft agar overlays whereby 100 µl of overnight D1614 culture and 100 µl phage stock were incubated 5 min at room temperature and mixed with 3 ml of 0.7% LB top agar. This molten mixture was poured on an LB plate and incubated at 30 ˚C for 18 hours. The top agar of plates showing confluent lysis was scraped into a 50 ml Falcon tube and 3 ml of modified SM was added for each scraped plate. This slurry was rocked at room temperature for 30 min, followed by centrifugation for 5 min at 10,000 x g to pellet debris. The resulting supernatant was filter-sterilized with a Millex-HA 0.45 µm syringe-driven filter (Millipore, Billerica, MA) and stored at 4 ˚C. Serial dilutions of phage stock into modified SM were used to determine host range and efficiency of plating on 27 *S. maltophilia* strains using the spot plate method detailed previously[164].

DLP5 stock for electron microscopy was prepared with LB agarose plates and soft agar. Plates with confluent lysis were overlaid with 3 ml of sterile MilliQ-filtered water, and the lysate

was centrifuged at 10,000 x g for 5 min, and filter sterilized with a 0.22 µm filter and immediately brought for observation. A 10 µl aliquot of the fresh DLP5 lysate was deposited onto the carbon-coated copper grid. After a 2 min incubation, the lysate was removed, and a 4% uranyl acetate solution was used to stain the grid for 30 s. A Philips/FEI (Morgagni) transmission electron microscope (TEM) with charge-coupled device camera at 80 kV (University of Alberta Department of Biological Sciences Advanced Microscopy Facility) was used to obtain TEM images. The capsid diameter, tail length and tail width of ten virions were measured using ImageJ and averages calculated using Microsoft Excel.

Phage DNA isolation, sequencing, and RFLP analysis

Genomic DNA was isolated from a high-titer DLP5 stock (10^9 PFU/ml). Lysate was clarified by spinning 10,000 x g for 10 min, and the supernatant was treated with 10 µl DNase I (Thermo Scientific, Waltham, MA), 100 µl 100x DNase I buffer (1 M Tris–HCl, 0.25 M $MgCl_2$, 10 mM $CaCl_2$), and 6 µl RNase (Thermo Scientific) and incubated 1 h at 37℃ to degrade contaminating bacterial nucleic acids. To this mixture, 400 µl of 0.5 M EDTA was added and SDS to a final concentration of 2 %. Proteinase K was then added to the solution for a final concentration of 400 µg/ml, followed by incubation at 55 ˚C overnight. A ½ volume of 6 M NaCl was then added to the solution and vortexed at high speed for 30 s followed by centrifugation at 17,900 x g for 30 min. The supernatant was transferred to a fresh tube with an equal volume of 100 % isopropanol and stored at -20 ˚C for at least 1 hour to overnight. The DNA was pelleted with centrifugation at 17,900 x g for 20 min at 4 ˚C followed by three 70 % ethanol washes. The pellet was dried at room temperature and resuspended in nuclease-free water. The purity and concentrations of eluted DNA were checked with a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA). A Nextera XT library was generated for paired-end sequencing on MiSeq (Illumina) platform using MiSeq v2 reagent kit, and the resulting reads were assembled using SPAdes 3.8.0[215].

Restriction fragment length polymorphism (RFLP) analysis was used with 36 FastDigest (Thermo Scientific) restriction enzymes: AccI, MspI, HpaII, HhaI, Bsh1236I, MauBI, PdmI, HaeIII, NheI, NdeI, AciI, EarI, SmaI, XbaI, BamHI, KpnI, ApaI, SfiI, PstI, SphI, HindIII, SacI, ClaI, SpeI, XhoI, PvuI, XmaI, EcoRV, DraI, StuI, HpaI, NotI, BgIII, NarI, SaII, and EcoRI. Restriction digest reactions were set up using the FastDigest system (Thermo Scientific) with 1

µl of an enzyme, 2 µl of buffer, 1 µg of phage DNA and nuclease-free water added to a final volume of 20 µl. Reactions were incubated at 37℃ for 20 min and separated on a 1% (wt/vol) agarose gel in 1x TAE (pH 8.0).

Determination of DLP5 lifestyle

Top agar overlay plates showing confluent lysis of D1614 by DLP5 were used to obtain resistant colonies. First, 3 ml of modified suspension media was added to the plate, and a sterile glass rod was used to gently skim the top of the agar. The SM was removed from the plates and placed in microcentrifuge tubes followed by centrifugation at 5,000 x g for 5 min. The supernatant was removed, and 1 ml of SM was added to the tube to resuspend the pellet. The tube was centrifuged again for 5 min at 5,000 x g, the supernatant removed, and 1 ml of SM added. This wash step was repeated three times in total. Following the final wash centrifugation, the supernatant was removed, and the pellet was resuspended in 500 µl LB broth. Serial dilution of the cells with LB, followed by plating onto LB plates, allowed for the isolation of single colonies following a 36-h incubation at 30 ˚C. A total of 22 colonies were picked into a 24-well plate with 1 ml of LB broth and incubated overnight with shaking at 30 ˚C. One well was left as an aseptic control and a single colony of wild type D1614 was inoculated into another well as a positive control. The next day, a 100 µl sample from each well was mixed with 1 ml of LB top agar impregnated with DLP5. The plate was then statically incubated at 30 ˚C overnight. The OD600 of the plate was obtained after 18 h incubation, with the results standardized to the negative control well. The wells with high OD600 values compared to wild type control were chosen for further analysis.

Superinfection experiments were performed using overnight cultures of the three potential lysogens and DLP5 at a $1x10^9$ PFU/mL titer in a top agar overlay assay. After a 16 h incubation at 30 ˚C, the plates were observed for plaque development. Lack of plaquing suggested DLP5 was stably maintained in the host D1614. Additionally, single colonies for each lysogen were used for colony PCR to detect the presence of DLP5. Overnight cultures of the D1614 lysogen grown in LB broth were used for plasmid extraction of DLP5 using both a bacterial artificial chromosome (BAC) extraction protocol and a Plasmid Mini Kit (Qiagen). The plasmid extraction protocol was performed on 2 ml of overnight culture centrifuged for 5 min at 6,000 x g. The pellet was resuspended in 100 µl resuspension buffer (50 mM glucose, 10 mM

EDTA, and 10 mM Tris-Cl pH 8.0). A 200 µl aliquot of lysis solution (0.2 M NaOH and 1% sodium dodecyl sulfate [SDS]) was added to the cell suspension which was mixed by inversion then incubated 5 min at room temperature. To the mixture, 150 µl 7.5 M ammonium acetate was added with 150 µl chloroform. The solution was mixed gently and chilled for 10 min on ice, followed by a 10 min 17,900 x g centrifugation step. The supernatant was then mixed with 200 µl precipitation solution (30 % polyethylene glycol 8000 and 1.5 M NaCl) and chilled on ice for 15 min. The solution was then centrifuged for 12 min at 14,000 x g, supernatant removed, and the DNA pellet was washed twice with 70 % ethanol. The pellet was dried 80 ˚C for 10 min then resuspended in 50 µl of water. Proteins were degraded by the addition of 5 µl 0.5 M EDTA (pH 8.0) with 1 µl proteinase K (25 mg/ml) and incubated at 55 ˚C for 2 hours. The solution was then mixed with 1 ml DNA binding resin and put though a Wizard DNA purification column (Promega). A 3 ml 80 % isopropanol wash, followed by a 10,000 x g 2 min centrifugation step, cleaned and dried the plasmid DNA. The DNA was then eluted from the column using 80 ˚C nuclease-free water applied to the column mesh, which was then incubated for 1 min followed by a 1 min 10,000 x g spin to elute the DNA.

Growth analysis of wild type D1614 versus the DLP5 lysogen

Single colony triplicate overnight cultures of *S. maltophilia* D1614 and the lysogen D1614::DLP5 were grown in LB broth at 30 ˚C with shaking. Subcultures (1:100) were performed for each sample using LB broth and grown to an OD600 of ~0.1 at 30 ˚C with 225 RPM shaking. Triplicate 200 µl aliquots of each subculture were distributed into 96 well plates with an LB broth control. The OD600 was then measured with a Wallac 1420 VICTOR2 multilabel counter (PerkinElmer, Waltham, MA) at the following time points: 0, 2, 4, 6, 8, and 10 h. The OD600 data were used to determine the growth rate (µ) with the established formula: $\log10\ N - \log10\ N_0 = (\mu/2.303)\ (t - t_0)$, whereby N_0 is the time zero (t_0) OD600 reading, and N is the final OD600 reading obtained at a specific time (t) in the experiment. Resulting data were analyzed with GraphPad Prism 7 (GraphPad Software Inc., San Diego, CA) to graph the growth curve and growth rate. Statistical analysis of the growth rate was performed in GraphPad Prism 7 using a two-way ANOVA with multiple comparisons. Additionally, cell counts were performed for each triplicate sample at every time point except at 6 h due to experimental error. A 10 µl aliquot was removed from each triplicate well and diluted into PBS. Three 10 µl aliquots from

each serial dilution were spotted onto LB plates and incubated 30 ˚C for 36 h to allow counting of single colonies.

Bioinformatic analysis of the DLP5 genome

A 96,542 bp contig assembled with SPAdes 3.8.0 was identified for further analysis. The mean coverage of the DLP5 genome is 80, and the Q40 is 99.7 %. Although no ambiguous regions in the contig were observed, amplification and sequencing of 15 randomly chosen sites were used to confirm the assembly. Open reading frames (ORFs) were identified with the GLIMMER plugin [172] for Geneious [217] using the Bacteria and Archaea setting, as well as GeneMarkS for phage [174]. Conserved domain searches were performed using CD-Search [175]. Phyre[218], HHblits[219,220], and I-TASSER[116] were used to gain insights into possible functions of hypothetical proteins when required. BLASTn and BLASTp (for full genomes and individual proteins, respectively) were used to gain information for each ORF and to identify any related phages[176]. BLASTp results above 1.00E-03 were not recorded, and the coding sequence (CDS) was annotated as hypothetical. Potential tRNAs were identified using the general tRNA model with tRNAscan-SE software[222]. Genomic comparisons were performed using the Large-Scale Genome Alignment Tool (LASTZ)[261,262] version 1.02.00, with DLP5 as the reference genome. The step length was set to 20, and the seed pattern was 12 of 19. Chaining and gapped alignment were performed, with both strands searched and both threshold scores, high-scoring segment pair and gapped, were set to 3,000.

RESULTS AND DISCUSSION

Isolation, morphology, host range, and RFLP analysis

Phage DLP5 (vB_SmaS_DLP_5) was isolated from Edmonton, Alberta, Canada garden soil using the *S. maltophilia* strain D1614. Transmission electron micrographs enable the classification of DLP5 as a B1 morphotype *Siphoviridae* phage[223] (Figure 5-1). The noncontractile tail of DLP5 averages 202.3 ± 4.0 nm long, and the head length and width measures 92.4 ± 2.6 and 84.5 ± 2.0 nm respectively. DLP5 has a narrow tropism within *S. maltophilia* clinical isolates, capable of successfully lysing 5/27 isolates tested. Infection by DLP5 results in clear plaques with defined borders averaging 0.5±0.2 mm and one step growth curves exhibit average burst sizes of 36. RFLP analysis suggests DLP5 DNA is modified as only

four of the 36 endonucleases were capable of digesting the DLP5 genome. The enzymes capable of digesting the genome target only GC sequences and have a four-base-pair (bp) target site: HaeIII; GGCC, HpaII; CCGG, AciI; CCGC, and Bsh1236I; CGCG. This result suggests the DNA contains A/T modifications, but the types and frequency of the modifications are unknown.

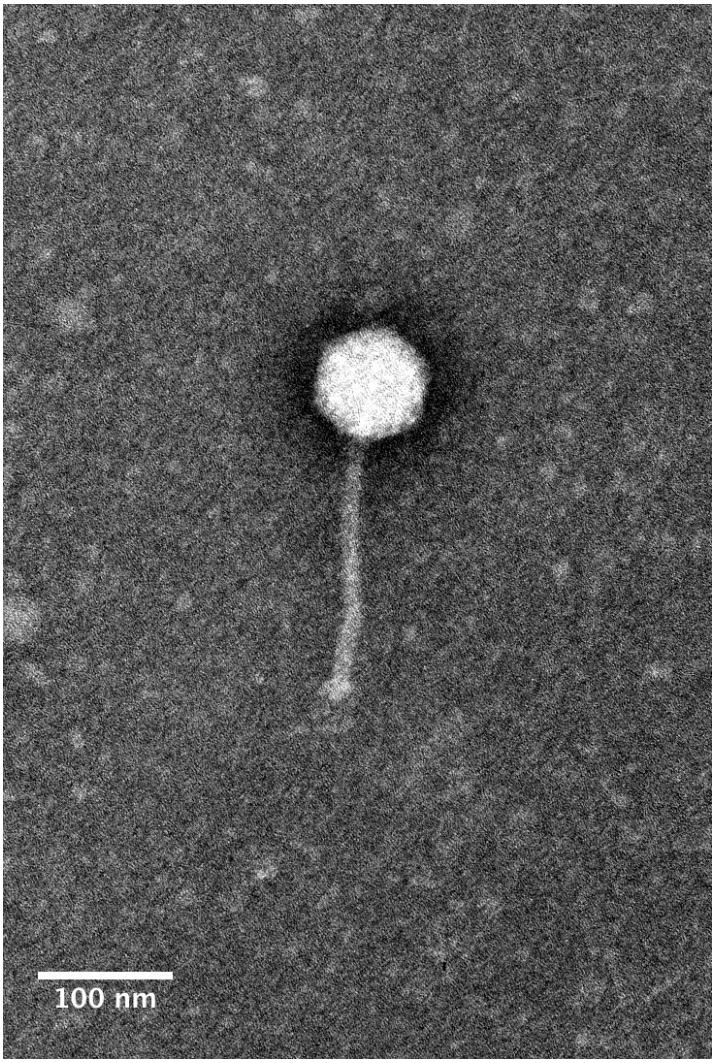

Figure 5-1: DLP5 *Siphoviridae* morphology. Phage lysate was applied to a carbon-coated copper grid, stained with 4 % uranyl acetate, and imaged at 180,000 x magnification by transmission electron microscopy. Average capsid length and width measurements based on nine virions are 92.4 ± 2.6 and 84.5 ± 2.0 nm respectively and the average tail length is 202.3 ± 4.0 nm.

Genomic characterization

The DLP5 genome was assembled into a 96,542 bp long contig, which can be accessed in GenBank under the accession MG189906. The GC content of DLP5 is 58.4 %, which is much lower than that of the host organisms' GC content of approximately 66 %. Inverted repeats of 346 bp were identified at the ends of the assembled contig. BLASTn analysis of DLP5 shows limited identity to other phages in the NCBI database, exhibiting maximum identity of 2% with *Xylella fastidiosa* phage Sano. Due to the limited results found with screening the NCBI database, the new genus *Delepquintavirus* was created with DLP5 as the type species. Characterization of DLP3, another *S. maltophilia Siphoviridae* phage isolated from Edmonton soil, led to the discovery of a second member of the *Delepquintavirus* genus. A comparison of these two phages is discussed below. Genomic characteristics of DLP5 include 149 coding domain sequences (CDS) and five tRNAs (Sup-CTA, Glu-TTC, Gly-TCC, Try-GTA, and Ser-GCT) (Table 5-1 and Figure 5-2). BLASTp analysis of the DLP5 proteins resulted in the identification of 74 hypothetical proteins, while 54 proteins did not have any identity to proteins in the NCBI virus database (Table 5-1). Conserved domain searches of the DLP5 proteins were completed to identify or support putative functions for each protein.

Conserved domain searches on the DLP5 proteins did provide some insight into functions for some of the hypothetical proteins (Table 5-2). For example, the portal protein (gp1) and serine protease XkdF (gp3) were identified using CD-Search, both of which were classified as hypothetical proteins with the BLASTp search. The TIP49 superfamily conserved domain was identified in two hypothetical proteins (gp41 and 134). This conserved region is present in RuvB-like proteins of eukaryotes and archaea[263], suggesting these proteins may be involved in homologous recombination. Other findings of interest from the CD-Search include two proteins (gp2 and 106) involved in chromosome partitioning with ParB domains. The gp106 protein also has a predicted SpoOJ superfamily domain. Other putative functions have been supported with the CD-Search results such as WecE (gp100), a protein involved in the synthesis of enterobacterial common antigen, a component of the bacterial outer membrane in members of the Enterobacteriaceae family[264]. This finding suggests DLP5 may be capable of altering the bacterial outer membrane of the host D1614 upon lysogenization, though this has not been confirmed experimentally.

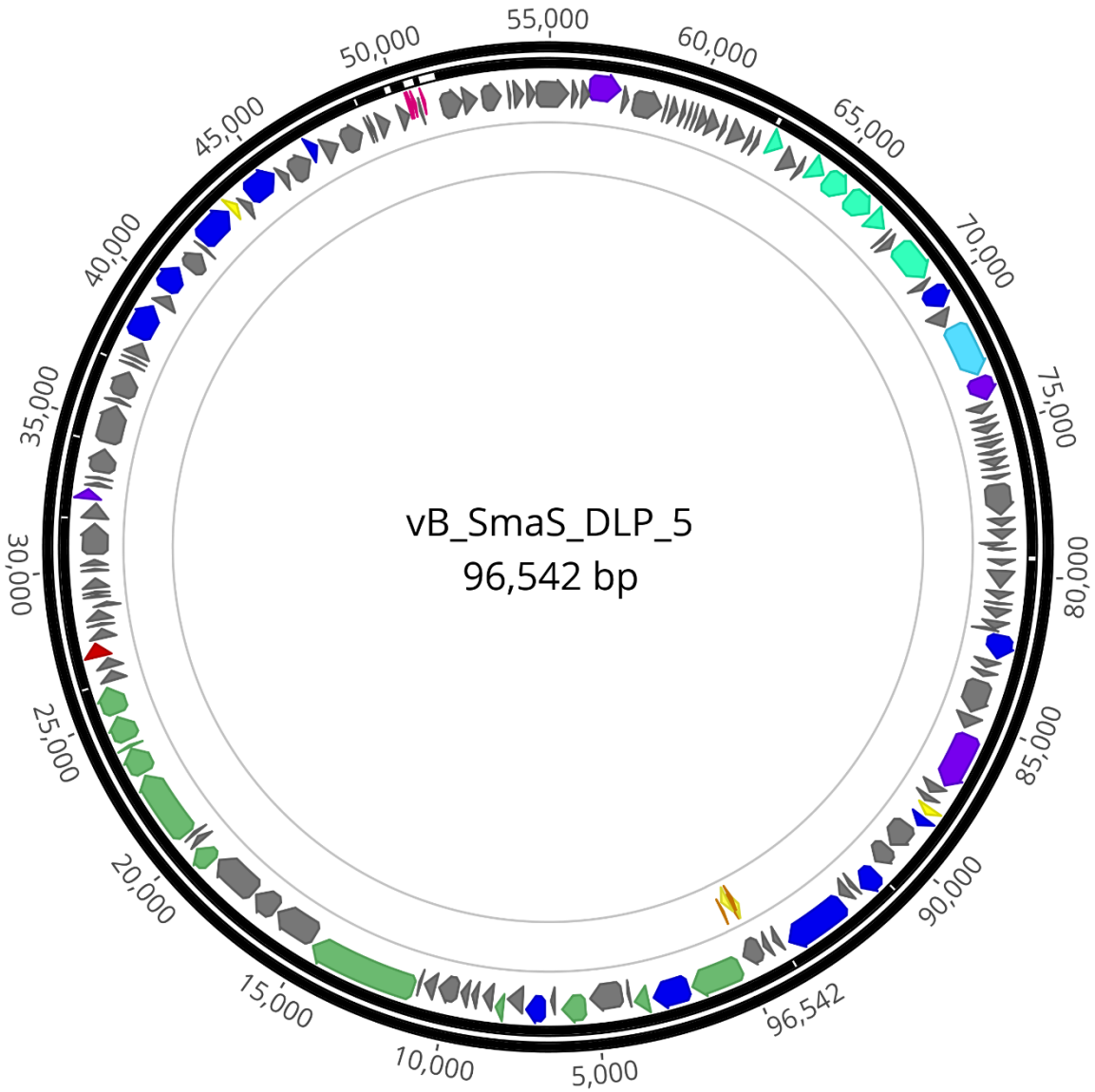

Figure 5-2: Genome map of DLP5. The inner circle of the genome map shows two repeat regions identified and represented as yellow and orange arrows. Predicted functions are grouped by color: teal; moron, grey; hypothetical, light blue; DNA packaging, pink; tRNA, red; lysis, green; virion morphogenesis, dark blue; DNA replication and repair, purple; auxiliary metabolism, and yellow; regulatory. The scale (bp) is shown on the outermost periphery of the map.

Table 5-1: Genome annotations for DLP5 from BLASTp and CD-Search data. Hits below 0.01 were not used and the function was annotated as hypothetical.

Gp	Interval	Length (AA)	Start codon	Putative function	Hit	Species	Cov. (%)	E value	ID (%)	Accession
1	78 - 1892	604	ATG	portal protein	hypothetical protein	*Burkholderia ubonensis*	95%	0.0E+00	49%	WP_060116606.1
2	1892 - 3232	446	ATG	ParB	hypothetical protein	*Burkholderia ubonensis*	61%	2.0E-74	45%	WP_060116607.1
3	3288 - 3890	200	ATG	serine protease XkdF	hypothetical protein	*Burkholderia ubonensis*	96%	6.0E-80	59%	WP_060116608.1
4	3883 - 4053	56	ATG	hypothetical protein	-					
5	4125 - 5351	408	GTG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	50%	5.0E-29	38%	WP_060116609.1
6	5387 - 6319	310	GTG	major capsid protein	hypothetical protein	*Burkholderia ubonensis*	99%	1.0E-119	57%	WP_060116610.1
7	6400 - 6669	89	ATG	hypothetical protein	-					
8	6741 - 7493	250	ATG	ribonuclease E	-					
9	7516 - 8127	203	GTG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	93%	1.0E-08	48%	WP_060116612.1
10	8127 - 8549	140	ATG	phage tail protein	hypothetical protein	*Burkholderia ubonensis*	99%	2.0E-36	50%	WP_060116613.1
11	8546 - 8995	149	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	93%	2.0E-16	29%	WP_060116614.1

12	9068 - 9361	97	ATG	hypothetical protein	-					
13	9372 - 9746	124	GTG	hypothetical protein	-					
14	9746 - 10501	251	ATG	hypothetical protein	hypothetical protein	*Nitrosomonas communis*	100 %	2.0E-91	57%	WP_046848997.1
15	10518 - 11012	164	ATG	hypothetical protein	hypothetical protein	*Pseudomonas aeruginosa*	74%	1.0E-03	29%	WP_073656054.1
16	11075 - 11197	40	TTG	hypothetical protein	-					
17	11203 - 14958	1251	ATG	tape measure protein	phage tail protein	*Pseudomonas aeruginosa*	20%	2.0E-37	44%	WP_077144342.1
18	14965 - 16548	527	ATG	hypothetical protein	hypothetical protein	*Lysobacter antibioticus*	100 %	4.0E-131	40%	ALN64563.1
19	16548 - 17525	325	ATG	hypothetical protein	hypothetical protein	*Lysobacter antibioticus*	98%	1.0E-44	36%	WP_057916382.1
20	17525 - 19201	558	ATG	hypothetical protein	hypothetical protein GLE_0033	*Lysobacter enzymogenes*	91%	2.0E-81	32%	ALN55392.1
21	19198 - 20016	272	ATG	minor tail protein	hypothetical protein	*Hahella ganghwensis*	100 %	8.0E-71	41%	WP_020406538.1
22	20016 - 20303	95	ATG	hypothetical protein	hypothetical protein	*Stenotrophomonas*	76%	1.0E-22	55%	WP_024957631.1
23	20300 - 20509	69	ATG	hypothetical protein	hypothetical protein MW7_0978	*Ralstonia* sp. PBA	94%	4.0E-09	45%	EIZ04531.1
24	20499 - 22979	826	GTG	phage tail protein	hypothetical protein COA75_03225	*Cellvibrionales*	86%	0.0E+00	42%	PCJ37747.1

25	22979 - 23977	332	ATG	tail assembly protein	tail assembly protein	*Xylella* phage Salvo	99%	3.0E-100	51%	AHB12242.1
26	23980 - 24156	58	ATG	tail assembly protein	tail assembly protein	*Xylella* phage Sano	100 %	2.0E-17	62%	AHB12066.1
27	24164 - 25135	323	ATG	tail assembly protein	tail assembly protein	*Xylella* phage Sano	100 %	0.0E+00	76%	AHB12065.1
28	25139 - 26155	338	ATG	tail protein	tail fiber protein	*Xylella* phage Sano	100 %	3.0E-145	63%	AHB12064.1
29	26232 - 26660	142	ATG	hypothetical protein	-					
30	26661 - 27149	162	ATG	hypothetical protein	-					
31	27146 - 27706	186	ATG	lysozyme	hypothetical protein	*Algiphilus aromaticivorans*	98%	3.0E-55	51%	WP_043767512.1
32	27708 - 28112	134	ATG	hypothetical protein	hypothetical protein	*Zunongwangia profunda*	42%	8.4E-01	33%	WP_013072096.1
33	28163 - 28333	56	ATG	hypothetical protein	-					
34	28353 - 28769	138	ATG	DUF2500 protein	hypothetical protein	*Rhizobium* sp. Root1220	26%	5.6E+00	39%	WP_056536871.1
35	28747 - 29028	93	GTG	hypothetical protein	-					
36	29128 - 29262	44	ATG	hypothetical protein	-					
37	29262 - 29426	54	ATG	hypothetical protein	-					

38	29426 - 29854	142	ATG	hypothetical protein	hypothetical protein AAY80_144	*Stenotrophomonas* phage DLP6	54%	4.0E-07	41%	AMQ66071.1
39	29958 - 30167	69	ATG	hypothetical protein	-					
40	30170 - 30472	100	ATG	hypothetical protein	-					
41	30532 - 31689	385	GTG	hypothetical protein	hypothetical protein	*Ochobactrum intermedium*	99%	1.0E-77	37%	WP_075042145.1
42	31794 - 32414	206	ATG	hypothetical protein	-					
43	32407 - 32859	150	ATG	phosphoglycerate kinase	-					
44	32859 - 33101	80	ATG	hypothetical protein	hypothetical protein	*Polaromonas naphthalenivorans*	93%	1.0E-08	36%	WP_011798018.1
45	33098 - 33286	62	ATG	hypothetical protein	-					
46	33347 - 34273	308	ATG	hypothetical protein	-					
47	34387 - 35778	463	ATG	hypothetical protein	-					
48	35778 - 36062	94	ATG	hypothetical protein	-					
49	36063 - 36983	306	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	88%	1.0E-27	33%	WP_060116641.1
50	37121 - 37333	70	ATG	hypothetical protein	-					

51	37323 - 37496	57	ATG	hypothetical protein	-					
52	37496 - 38056	186	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	93%	1.0E-40	38%	WP_060116642.1
53	38053 - 39342	429	GTG	DnaB helicase	hypothetical protein	*Nitratireductor aquibiodomus*	98%	6.0E-82	37%	WP_083212884.1
54	39403 - 39912	169	ATG	hypothetical protein	-					
55	39905 - 40870	321	GTG	DnaG primase	hypothetical protein	*Burkholderia ubonensis*	90%	9.0E-70	44%	WP_060116518.1
56	40948 - 41721	257	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	87%	4.0E-71	48%	WP_060116519.1
57	41757 - 41984	75	ATG	hypothetical protein	putative DNA primase	*Pseudomonas* phage PAE1	48%	3.0E-03	56%	YP_009215749.1
58	41981 - 43345	454	TTG	Superfamily II DNA or RNA helicase	hypothetical protein	*Burkholderia ubonensis*	97%	2.0E-125	47%	WP_060116520.1
59	43342 - 43743	133	ATG	transcriptional regulator	hypothetical protein	*Burkholderia ubonensis*	97%	2.0E-18	36%	WP_060116521.1
60	43746 - 44117	123	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	95%	8.0E-13	38%	WP_060116522.1
61	44117 - 45319	400	ATG	RecA	hypothetical protein	*Burkholderia ubonensis*	90%	3.0E-157	61%	WP_080421533.1
62	45319 - 45693	124	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	96%	7.0E-33	50%	WP_060116524.1
63	45690 - 46574	294	ATG	hypothetical protein	hypothetical protein WK13_34410	*Burkholderia ubonensis*	98%	2.0E-97	51%	KVR21633.1

64	46571 - 47068	165	ATG	RuvC	hypothetical protein	*Burkholderia ubonensis*	86%	2.0E-29	38%	WP_060116526.1
65	47046 - 47678	210	ATG	hypothetical protein	MazG	*Bacteriovorax* sp. BAL6_X	41%	9.0E-03	37%	WP_021266999.1
66	47761 - 48564	267	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	93%	3.0E-90	57%	WP_060116528.1
67	48749 - 48907	52	ATG	hypothetical protein	-					
68	48912 - 49025	37	ATG	hypothetical protein	-					
69	49108 - 49584	158	ATG	hypothetical protein	-					
70	49873 - 50262	129	ATG	hypothetical protein	hypothetical protein BN7874_027	Phage NCTB	63%	6.0E-17	50%	SBV38196.1
71	50616 - 50741*	41	ATG	hypothetical protein	-					
72	51339 - 52109	256	ATG	hypothetical protein	hypothetical protein	*Achomobacter* sp. KCJK1731	15%	1.0E-05	56%	WP_063959466.1
73	52106 - 52672	188	GTG	hypothetical protein	hypothetical protein	*Burkholderia gladioli*	38%	2.0E-22	58%	WP_013691858.1
74	52692 - 53489	265	GTG	hypothetical protein	SPFH domain-containing protein	*Stenotrophomonas panacihumi*	98%	4.0E-140	73%	WP_057645969.1
75	53585 - 53812	75	ATG	hypothetical protein	-					
76	53809 - 54249	146	GTG	hypothetical protein	-					

77	54246 - 54596	116	GTG	hypothetical protein	hypothetical protein	*Paenibacillus dendritiformis*	90%	8.0E-27	47%	WP_006675361.1
78	54589 - 55743	384	GTG	hypothetical protein	-					
79	55743 - 56084	113	ATG	hypothetical protein	hypothetical protein	*Cupriavidus* sp. USMAA1020	95%	5.0E-05	33%	WP_071070001.1
80	56097 - 56438	113	ATG	hypothetical protein	-					
81	56363 - 57469	368	GTG	UDP-glucose 4-epimerase	UDP-glucose 4-epimerase GalE	*Hydrogenibacillus schlegelii*	91%	2.0E-52	37%	WP_066198487.1
82	57504 - 57824	106	ATG	hypothetical protein	hypothetical protein NOXIFER_222	*Pseudomonas* phage Noxifer	78%	3.0E-11	38%	ARV77387.1
83	57890 - 58948	352	ATG	hypothetical protein	hypothetical protein	*Rhizobium laguerreae*	61%	7.0E-04	23%	WP_077979619.1
84	58941 - 59162	73	ATG	hypothetical protein	hypothetical protein	*Stenotrophomonas maltophilia*	91%	4.0E-11	42%	WP_014036685.1
85	59159 - 59530	123	ATG	hypothetical protein	unnamed protein product	*Ralstonia* phage phiRSL1	24%	1.3E-02	63%	YP_001950148.1
86	59527 - 59781	84	ATG	hypothetical protein	-					
87	59778 - 59954	58	ATG	hypothetical protein	-					
88	59963 - 60145	60	ATG	hypothetical protein	hypothetical protein IME11_78	*Escherichia* phage IME11	93%	3.0E-06	39%	YP_006990683.1
89	60148 - 60513	121	ATG	hypothetical protein	-					

90	60510 - 60977	155	ATG	hypothetical protein	-					
91	60974 - 61249	91	ATG	hypothetical protein	-					
92	61278 - 61922	214	ATG	hypothetical protein	-					
93	61929 - 62207	92	ATG	hypothetical protein	-					
94	62204 - 62443	79	GTG	hypothetical protein	-					
95	62627 - 63247	206	ATG	Rhomboid membrane protein	hypothetical protein CBE50_03270	*Flammeovirgaceae* sp. TMED290	60%	3.0E-10	28%	OUX56367.1
96	63306 - 63899	197	ATG	hypothetical protein	hypothetical protein	*Pseudomonas* sp. GM84	82%	6.0E-04	27%	WP_008100375.1
97	63899 - 64156	85	ATG	hypothetical protein	-					
98	64232 - 64870	212	ATG	PIG-L family deacetylase	hypothetical protein A3J57_00650	*Candidatus Wildermuthbacteria*	96%	1.0E-23	34%	OHA69194.1
99	64875 - 65816	313	ATG	WcaG	WcaG	*Candidatus Kaiserbacteria*	95%	3.0E-50	34%	KKW06749.1
100	65816 - 66862	348	ATG	WecE	hypothetical protein	*Lachnospiraceae* sp. M18-1	95%	4.0E-51	35%	WP_016298419.1
101	66859 - 67539	226	GTG	methyltransferase	hypothetical protein A3D64_01355	*Candidatus Wildermuthbacteria*	76%	2.0E-15	29%	OHA70649.1
102	67595 - 67795	66	ATG	hypothetical protein	-					

103	67792 - 68214	140	ATG	hypothetical protein	-					
104	68211 - 69665	484	ATG	glycosyltransferase	hypothetical protein	*Ochobactrum intermedium*	76%	7.0E-28	25%	WP_075042103.1
105	69662 - 69973	103	GTG	hypothetical protein	-					
106	69945 - 70799	284	GTG	ParBc	hypothetical protein	*Ochobactrum intermedium*	97%	2.0E-81	45%	WP_075042105.1
107	70799 - 71446	215	ATG	hypothetical protein	-					
108	71409 - 73322	637	GTG	large terminase subunit	hypothetical protein	*Nitratireductor aquibiodomus*	86%	0.0E+00	56%	WP_065815631.1
109	73334 - 74227	297	GTG	hypothetical protein	hypothetical protein	*Rhizobium* sp. Leaf383	100 %	2.0E-71	46%	WP_062600990.1
110	74224 - 74634	136	ATG	DUF3310 protein	DUF3310 protein	*Mycobacterium abscessus*	81%	4.0E-17	38%	WP_074322358.1
111	74696 - 74938	80	ATG	hypothetical protein	expressed ef-hand protein	*Micromonas commoda*	63%	1.5E+00	33%	XP_002499896.1
112	74940 - 75332	130	ATG	hypothetical protein	hypothetical protein	*Streptomyces rimosus*	70%	4.9E-02	27%	WP_050503634.1
113	75401 - 75499	32	ATG	hypothetical protein	-					
114	75496 - 75813	105	ATG	hypothetical protein	hypothetical protein	*Phaeospirillum molischianum*	80%	2.0E-14	42%	WP_002731271.1
115	75810 - 76064	84	ATG	hypothetical protein	hypothetical protein	*Klebsiella pneumoniae*	52%	3.4E+00	39%	WP_087847212.1

116	76054 - 76428	124	ATG	hypothetical protein	-					
117	76419 - 76571	50	ATG	hypothetical protein	-					
118	76654 - 76905	83	GTG	hypothetical protein	-					
119	76961 - 78133	390	ATG	hypothetical protein	hypothetical protein	*Sphingomonas* sp. Root50	63%	7.0E-36	44%	WP_056601442.1
120	78133 - 78483	116	ATG	hypothetical protein	hypothetical protein	*Sphingomonas*	97%	8.0E-41	56%	WP_056363031.1
121	78509 - 78970	153	ATG	hypothetical protein	hypothetical protein	*Leptospira* sp. P2653	81%	1.0E-34	48%	WP_020782964.1
122	78936 - 79166	76	ATG	hypothetical protein	-					
123	79163 - 79342	59	ATG	hypothetical protein	-					
124	79534 - 79842	102	ATG	hypothetical protein	-					
125	79868 - 80602	244	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	87%	2.0E-23	33%	WP_060116587.1
126	80599 - 80988	129	ATG	hypothetical protein	hypothetical protein L522_1762	*Bordetella bronchiseptica* MBORD707	37%	9.0E-19	73%	KDD09863.1
127	80981 - 81163	60	ATG	hypothetical protein	hypothetical protein PBI_121Q_495	*Escherichia* phage 121Q	85%	3.0E-11	53%	YP_009102082.1
128	81163 - 81636	157	ATG	DUF1643 protein	DUF1643 protein	*Burkholderia pseudomultivorans*	100 %	4.0E-45	52%	WP_059518962.1

129	81633 - 81974	113	GTG	hypothetical protein	hypothetical protein Sano_18	*Xylella* phage Sano	83%	2.0E-26	53%	AHB12038.1
130	81954 - 82067*	37	GTG	hypothetical protein	-					
131	82047 - 82943	298	ATG	DNA ligase	ATP-dependent DNA ligase	*Escherichia* phage CAjan	97%	2.0E-86	45%	YP_009196846.1
132	82940 - 83356	138	GTG	hypothetical protein	-					
133	83356 - 83649	97	ATG	hypothetical protein	-					
134	83709 - 84920	403	ATG	hypothetical protein	hypothetical protein	*Ochobactrum intermedium*	83%	2.0E-77	40%	WP_075042145.1
135	84917 - 85489	190	ATG	hypothetical protein	-					
136	85583 - 87520	645	ATG	pyruvate phosphate dikinase	pyruvate, phosphate dikinase	*Burkholderia ubonensis*	99%	0.0E+00	50%	WP_060116597.1
137	87566 - 87982	138	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	83%	2.0E-08	30%	WP_060116598.1
138	87984 - 88325	113	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	97%	2.0E-12	38%	WP_060116599.1
139	88294 - 88692	132	ATG	transcriptional repressor	transcriptional repressor	*Pseudomonas stutzeri*	71%	1.1E+00	31%	WP_015278930.1
140	88682 - 89152	156	GTG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	100 %	1.0E-40	50%	WP_060116600.1
141	89194 - 90210	338	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	99%	4.0E-90	43%	WP_060116601.1

142	90207 - 90995	262	ATG	tyrosine phosphatase protein	hypothetical protein	*Burkholderia ubonensis*	88%	1.0E-44	39%	WP_060116602.1
143	91099 - 91932	277	GTG	hypothetical protein	hypothetical protein	*Nitratireductor aquibiodomus*	47%	7.0E-06	32%	WP_065815622.1
144	91922 - 92176	84	ATG	hypothetical protein	pleiotropic drug resistance protein ABC superfamily	*Phytophthora sojae*	73%	2.0E+00	37%	XP_009514447.1
145	92173 - 92655	160	ATG	hypothetical protein	hypothetical protein	*Burkholderia ubonensis*	70%	1.2E-02	26%	WP_060116604.1
146	92618 - 94903	761	ATG	DNA polymerase I	hypothetical protein	*Burkholderia ubonensis*	98%	0.0E+00	57%	WP_060116605.1
147	95076 - 95336	86	ATG	hypothetical protein	-					
148	95447 - 95728	93	ATG	hypothetical protein	-					
149	95725 - 96516*	262	GTG	hypothetical protein	hypothetical protein	*Acinetobacter* sp. TGL-Y2	70%	3.0E-17	44%	WP_067663791.1

Gp: gene product

Table 5-2: Conserved domain results from a CD-Search using 149 predicted DLP5 proteins.

Gene product	Hit type	PSSM-ID	Interval	E-Value	Accession	Short name	Superfamily
1	superfamily	327517	45 - 472	2.56E-55	cl19194	Phage_portal superfamily	-
2	specific	214678	351 - 443	1.73E-05	smart00470	ParB	cl02129
3	superfamily	317012	4 - 113	2.44E-24	cl24270	Peptidase_S78_2 superfamily	-
6	superfamily	331903	20 - 307	1.68E-12	cl27082	Phage_capsid superfamily	-
8	superfamily	331378	7 - 111	0.0018	cl26557	RNase_E_G superfamily	-

10	superfamily	321796	6 - 136	3.41E-13	cl02089	Phage_tail_S superfamily	-
17	superfamily	331332	521 - 923	8.02E-16	cl26511	Neuromodulin_N superfamily	-
17	superfamily	333387	161 - 228	0.0046	cl28567	HI1514 superfamily	-
21	specific	312753	188 - 264	7.20E-19	pfam09356	Phage_BR0599	cl10710
21	superfamily	331404	18 - 261	1.04E-13	cl26583	DUF2163 superfamily	-
24	specific	316107	207 - 368	9.32E-13	pfam13550	Phage-tail_3	cl26145
28	specific	238058	96 - 182	5.84E-09	cd00110	LamG	cl22861
31	superfamily	331815	3 - 186	8.57E-31	cl26994	Glyco_hydro_108 superfamily	-
34	superfamily	313826	22 - 95	0.0095	cl11292	DUF2500 superfamily	-
41	superfamily	332389	76 - 371	1.84E-40	cl27568	TIP49 superfamily	-
53	superfamily	333705	167 - 405	1.57E-23	cl28885	RecA-like_NTPases superfamily	-
55	superfamily	331610	31 - 312	1.16E-10	cl26789	Toprim_N superfamily	-
58	superfamily	331760	51 - 428	1.34E-37	cl26939	DEXDc superfamily	-
59	superfamily	322007	56 - 110	0.0001	cl02600	HTH_MerR-SF superfamily	-
61	superfamily	333705	72 - 267	1.66E-49	cl28885	RecA-like_NTPases superfamily	-
64	superfamily	328743	1 - 140	9.69E-11	cl21482	RuvC_resolvase superfamily	-
66	superfamily	328734	49 - 110	0.0079	cl21469	HDc superfamily	-
74	specific	307341	24 - 208	7.49E-17	pfam01145	Band_7	cl19107
76	superfamily	330398	87 - 125	0.0081	cl25577	PKS_ER superfamily	-
81	superfamily	330230	23 - 351	8.80E-77	cl25409	SDR superfamily	-
95	superfamily	328780	60 - 190	1.31E-12	cl21536	Rhomboid superfamily	-
98	specific	308281	5 - 122	9.84E-13	pfam02585	PIG-L	cl00929

99	specific	223528	1 - 297	1.35E-37	COG0451	WcaG	cl25660
100	specific	223476	14 - 341	3.00E-55	COG0399	WecE	cl18945
101	superfamily	327401	34 - 135	0.0019	cl17173	AdoMet_MTases superfamily	-
104	specific	223515	1 - 323	0.0011	COG0438	RfaB	cl28208
106	specific	308032	18 - 105	8.38E-17	pfam02195	ParBc	cl02129
106	specific	224392	34 - 187	1.50E-07	COG1475	Spo0J	cl26722
109	superfamily	332234	161 - 196	0.0090	cl27413	Thy1 superfamily	-
110	specific	314594	17 - 68	3.78E-13	pfam11753	DUF3310	cl13237
128	specific	311648	12 - 143	5.47E-44	pfam07799	DUF1643	cl01787
131	superfamily	325160	25 - 186	1.34E-18	cl12015	Adenylation_DNA_ligase_like superfamily	-
131	superfamily	330238	110 - 289	6.46E-07	cl25417	CDC9 superfamily	-
134	superfamily	332389	206 - 341	3.02E-34	cl27568	TIP49 superfamily	-
136	superfamily	331842	16 - 471	0	cl27021	PtsP superfamily	-
139	superfamily	333066	91 - 122	5.68E-05	cl28246	DnaJ superfamily	-
140	superfamily	330819	44 - 125	3.25E-08	cl25998	CDC14 superfamily	-
142	superfamily	330819	111 - 185	9.42E-06	cl25998	CDC14 superfamily	-
146	superfamily	322025	134 - 731	6.84E-79	cl02626	DNA_pol_A superfamily	-

DLP3 relatedness to DLP5

While characterizing the *Siphoviridae* phage DLP3, it was evident though BLASTn, and BLASTp searches that these two phages are closely related. The genome size of DLP3 and DLP5 are similar (96,852 versus 96,542 bp), as is their GC content (58.3 versus 58.4 %) (Table 5-3). BLASTn analysis of DLP3 against DLP5 revealed 81 % identity over 90% of the DLP5 genome (0.0 E-value). A LASTZ alignment of the phages using DLP5 as the reference sequence shows sequence identity > 30 % between their genomes, represented as a mustard yellow color in the consensus. BLASTn alignment between DLP3 to DLP5 gives a query coverage of 89 % and 81 % identity. Two small stretches with a breakdown in identity are observed between the DLP5 and DLP3 genomes which is viewed as a breakdown in the alignment blocks of the dot plot around 30,000 and 60,000 bp (Figure 5-3). These stretches correspond to four genes (DLP05_037 to DLP05_039, and DLP05_087) encoding hypothetical proteins with no significant matches in the NCBI database, no conserved domains, and no significant results when using structural prediction software such as HHpred and Phyre. Two additional genes, DLP05_045 and DLP05_068, which encode hypothetical proteins, are not present in the DLP3 genome. The DLP05_045 gene product shares 93 % coverage with 36 % identity (1.0E-08) to a *Polaromonas naphthalenivorans* hypothetical protein which contains no conserved domains, while the DLP05_068 gene product did not have similar sequences in the NCBI database or conserved domains predicted. Both phages encode five tRNAs, with four of the five tRNAs sharing the same specificity: Sup-CTA, Glu-TTC, Ser-GCT, Tyr-GTA. DLP3 differs from DLP5 with respect to the fifth tRNA, which is Ile-GAT in DLP3 but Gly-TCC in DLP5. The amino acid usage of each phage does not explain the differences, as they each have the same usage rates for isoleucine (4.7 %) and glycine (7.7 %), but there are several nucleotide changes observed for this region when comparing DLP3 to DLP5.

Table 5-3: Comparison of the DLP3 and DLP5 genomic traits.

Phage	Size (bp)	GC content (%)	ORFs	tRNAs
DLP3	96,852	58.3	148	5
DLP5	96,542	58.4	149	5

Besides the genomic similarities observed between DLP3 and DLP5, they also share morphological similarities. The DLP3 and DLP5 phages have the same measurement averages

for their head width (84 nm) and length (92 nm), as well as tail length (202 nm). The host range of DLP3 versus DLP5 is significantly different, with DLP3 exhibiting a broad host range infecting 21 *S. maltophilia* strains, while DLP5 only infects five (Table 1-2). Of the five strains DLP5 is capable of infecting (102, 249, D1571, D1614, and D1576), DLP3 is only unable to infect the DLP5 host strain D1614. The virion morphogenesis proteins of DLP3 range in sequence identity to the DLP5 equivalents from 43.8 % (gp28) to 93.2 % (gp6), with all but gp28 having sequence identity greater than 70 %. The gp28 proteins of DLP3 and DLP5 are related to the *Xylella* phage Sano tail fiber protein. The annotation of tail fiber for the Sano protein was based solely on synteny to the *Burkholderia* phage BcepNazgul[255]. There are no shared conserved domains identified between these two proteins, and the BcepNazgul protein is 1,270 amino acids long while the DLP3 and DLP5 gp28 proteins and the Sano tail fiber protein are only around 340 amino acids long; thus, it appears assigning tail fiber to the Sano phage protein was incorrect.

The gp28 proteins of DLP3 and DLP5 may play a role in host range due to the variability observed between these proteins. The DLP5 gp28 protein shares 62.5 % identity with the Sano tail fiber, while DLP3 only has 39.6 % identity to the Sano protein, and 41.8 % identity to the DLP5 protein. The first 242 N-terminal amino acids of the DLP3 and DLP5 gp28 consensus show a high degree of variability at 24.5 % pairwise identity. The remaining 104 amino acids from 243 to the C-terminal have high pairwise identity at 88.5 %. Variability in the N-terminal region was also observed with a MUSCLE alignment using all three proteins from DLP3, DLP5, and Sano (Figure 5-4). The first 258 N-terminal amino acids of the protein consensus share only 35.4 % pairwise identity, which increases to 77.8 % over the remaining 104 amino acids. The Figure 5-4 alignment shows at least 9 insertion/deletion events have occurred within the DLP3 gene resulting in gaps and insertions in the translated DLP3 protein compared to the Sano and DLP5 proteins. Further investigation into gp28 of DLP3 and DLP5 may help elucidate if this protein plays a role in the host range of the *Delepquintavirus* phages.

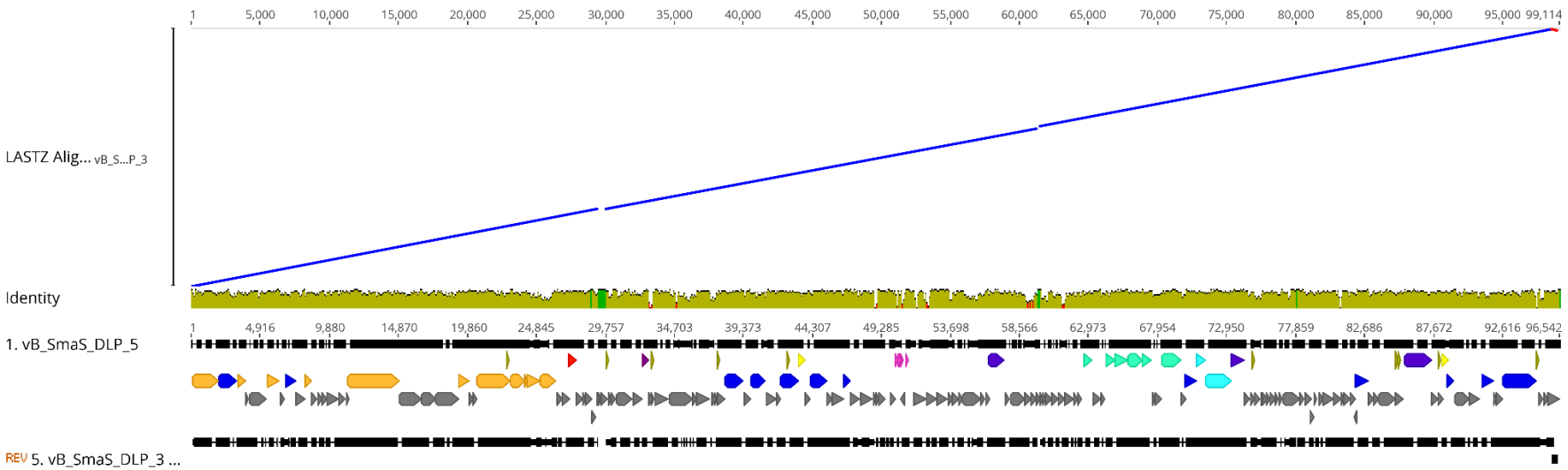

Figure 5-3: Genomic alignment of DLP3 to DLP5 using the Large-Scale Genome Alignment Tool. Identity is indicated by color: mustard yellow; > 30 %, green; deletion in DLP3, red; deletion in DLP5.

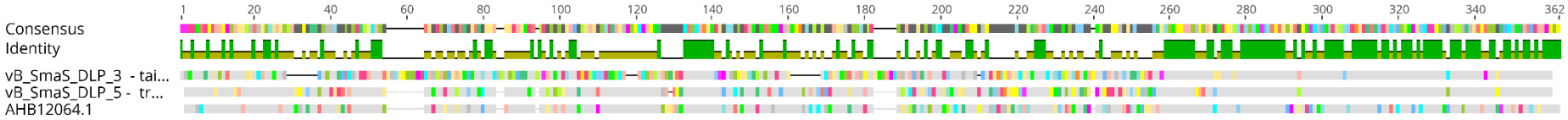

Figure 5-4: MUSCLE multiple sequence alignment of gp28 proteins from DLP3 and DLP5, and tail fiber protein of *Xylella* phage Sano. There appears to be some amino acid conservation towards the C-terminal of the proteins versus the N-terminal when observing the consensus identity of all three proteins.

Lysogenic conversion of *S. maltophilia* D1614 by bacteriophage DLP5

Stable lysogens of D1614::DLP5 were isolated for further study. Due to the presence of ParB conserved domains within two of the DLP5 proteins, it was hypothesized DLP5 was stably maintained as a phagemid within the lysogen. Plasmid preps of the lysogen run on a 1 % agarose gel show a band tracking at the same height as the DLP5 gDNA prepped from phage lysate, confirming DLP5 lysogenizes its host as a phagemid (Figure 5-5).

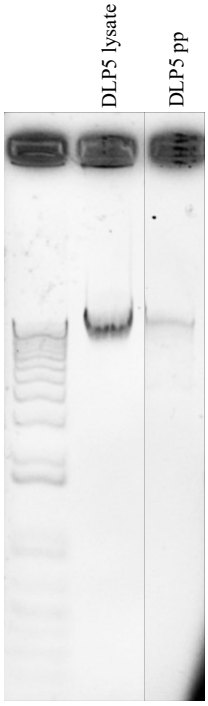

Figure 5-5: DLP5 lysogenizes as a phagemid in *S. maltophilia* D1614. 1 % agarose gel stained with ethidium bromide and gel image spliced to show applicable data. Lane 1: 1 Kbp Plus DNA ladder (Invitrogen), lane 2: DLP5 gDNA from phage lysate, lane 3: DLP5 plasmid prepped from D1614::DLP5 lysogen.

While working with the wild type D1614 and lysogen D1614::DLP5 cultures, it became apparent there were changes in the growth rate of the lysogen compared to the wild type host. A growth curve analysis was completed with cell forming unit counts to determine what effect DLP5 lysogenization has on the growth rate of the host (Figure 5-7 and Table 5-4). The results show there is an initial increase in growth rate observed over the zero to two hour time interval at a statistically significant level ($P < 0.001$), much like the D1571::DLP3 lysogen discussed in Chapter 3 (Figure 5-6 and Figure 5-7). An increased growth rate is then observed for the D1614

wild type strain over the lysogen from two to four, four to six and six to eight-hour time intervals at a statistically significant level ($P<0.0001$, 0.01 and 0.01 respectively). The change in growth rate from zero to two and two to four hours could suggest the lysogen is growing well during the first two hours, but after two hours the prophage switches to the lytic cycle and lyses the host cells. To account for this possibility colony forming unit (CFU) counts were obtained from the growth curve experiments (Table 5-4). The results show DLP5 is not lysing the host cells but is stably maintained in the host as there is no decrease in the CFUs of the lysogen; thus, DLP5 must not be lysing the host cells. The reason for the decreased growth rate of the lysogen between two and eight hours is unknown.

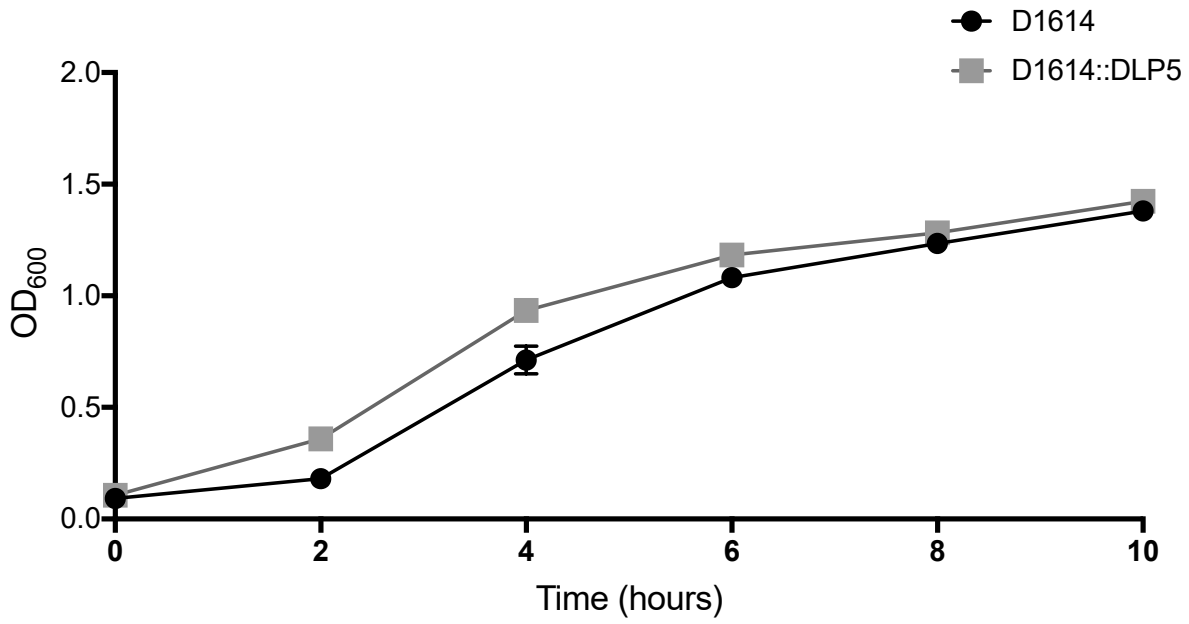

Figure 5-6: Growth curve analysis of wild type D1614 and lysogen D1614::DLP5 grown over ten hours. Results from biological and mechanical triplicate experiments were averaged and the mean plotted with the standard deviation represented as error bars.

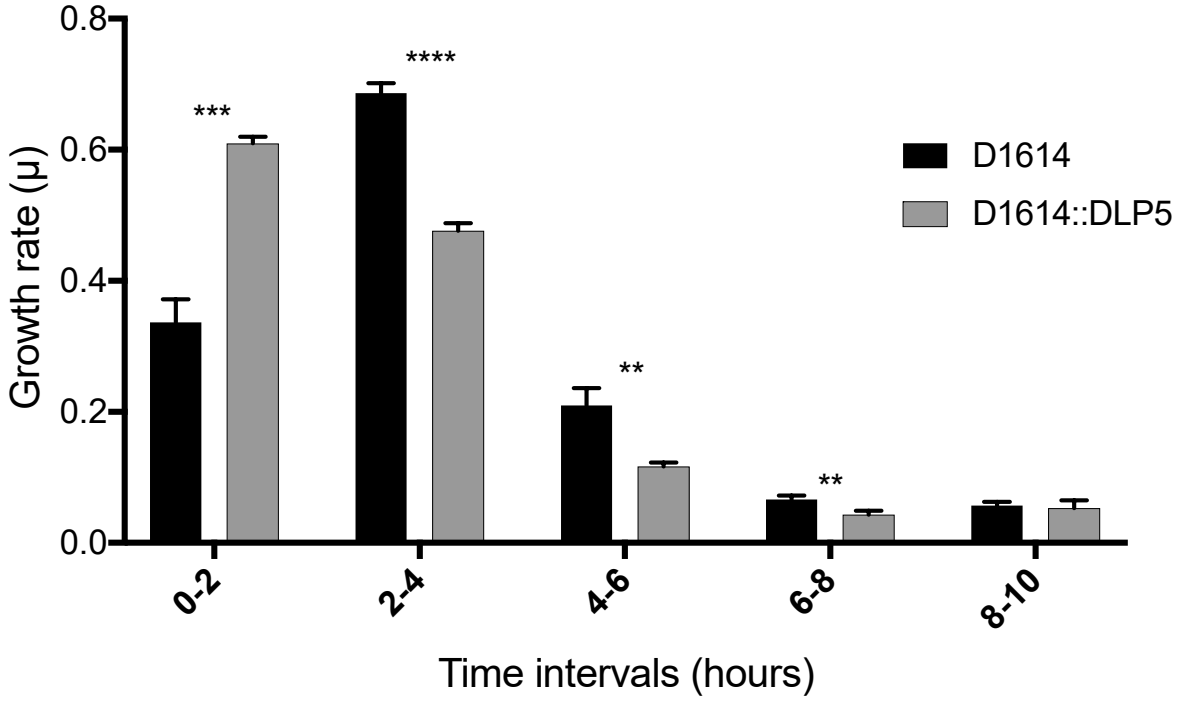

Figure 5-7: Growth rate of wild type D1614 versus lysogen D1614::DLP5. Data obtained from the OD600 growth curve measurements was converted to growth rate and the mean plotted. Error bars represent standard deviation. Two-way ANOVA with multiple comparisons indicates statistical significance (P value): **; <0.01, ***; <0.001, ****; <0.0001.

Table 5-4: Colony forming units (CFU) obtained from the D1614 and D1614::DLP5 growth curve experiment. Mechanical and biological triplicates were performed for each bacterial strain at the time points listed. The 6 h time point was excluded due to experimental error.

Time (hours)	Bacterial strains (CFU ± SD)	
	D1614	D1614::DLP5
0	$7.6 \times 10^7 \pm 2.2$	$2.4 \times 10^8 \pm 0.7$
2	$1.5 \times 10^8 \pm 0.5$	$1.3 \times 10^9 \pm 0.5$
4	$9.8 \times 10^8 \pm 3.7$	$2.0 \times 10^9 \pm 1.0$
8	$6.1 \times 10^9 \pm 3.6$	$9.3 \times 10^9 \pm 3.8$
10	$5.9 \times 10^9 \pm 0.5$	$1.2 \times 10^{10} \pm 0.5$

CONCLUSION

Isolation and characterization of the novel *Siphoviridae* phage DLP5 led to the creation of a new genus of phages called *Delepquintavirus*. Sequencing of DLP5 led to the assembly of a 96,542 bp contig encoding 149 open reading frames with 58.4 % GC content. DLP5 phage is capable of lysogenizing *S. maltophilia* D1614 as a phagemid, potentially using two proteins containing ParB conserved domains for partitioning of the phagemid into daughter cells of the dividing lysogen. Lysogenization of D1614 by DLP5 leads to a statistically significant increase in growth rate over the first two hours; though, this trend disappears over the remaining growth periods with the wild type strain growing faster than the lysogen. Lysogenic conversion of *S. maltophilia* strains by temperate phages though increased growth rate, increased antibiotic resistance, or though the contribution of virulence factors has been noted with many characterized *Stenotrophomonas* phages[154,156,159,161,166]. The findings presented here highlight the importance of characterizing both temperate and virulent phages to better understand the dynamic role of phages in the evolution of *S. maltophilia* strains.

ACKNOWLEDGEMENTS

The authors thank David Speert of the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR) for the kind gift of numerous *S. maltophilia* strains. The authors would also thank the Provincial Laboratory for Public Health— North (Microbiology), Alberta Health Services, for access to *S. maltophilia* clinical isolates. The authors acknowledge helpful scientific discussions with other members of the Dennis lab. This research was supported by Discovery Grant #238414 to J.J.D. from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Chapter 6

The Isolation and Characterization of *Stenotrophomonas maltophilia* T4-like Bacteriophage DLP6

Portions of this chapter have been submitted as:

Peters, D. L., P. Stothard and J. J. Dennis. 2017. The Isolation and Characterization of *Stenotrophomonas maltophilia* T4-like Bacteriophage DLP6. PLoS ONE. **12**: e0173341.

OBJECTIVES

The objectives of this research were to characterize the T4-like bacteriophage DLP6 by sequencing and annotating the genome then phylogenetically compare it to other members of the T4-superfamily.

MATERIALS AND METHODS

Bacterial strains and growth conditions

Five *S. maltophilia* strains were acquired from the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR; Vancouver, BC). The *S. maltophilia* strains used for phage isolation from soil samples were D1585, D1571, D1614, D1576, and D1568. An additional 22 *S. maltophilia* strains were gifted from the Provincial Laboratory for Public Health—North (Microbiology), Alberta Health Services, for host range analysis. All strains were grown aerobically overnight at 30℃ on half-strength Luria-Bertani (½ LB) solid medium or in ½ LB broth with shaking at 225 RPM.

Phage isolation, propagation, host range analysis, and electron microscopy

DLP6 was isolated from planter soil located at the Kinsman Sports Center in Edmonton, Alberta, Canada using strain D1571 and a previously described extraction protocol[163]. Propagation of DLP6 was performed using soft agar overlays: 100 µl liquid culture and 100 µl phage stock were incubated 20 min at room temperature, mixed with 3 ml 0.7% 1⁄2 LB top agar, overlaid on a plate of ½ LB solid medium, and incubated at 30℃ until plaque formation was complete. High titer stocks were made by overlaying plates showing confluent lysis with 3 ml modified SM; then the top agar was scraped into a sterile Falcon tube. 5-min centrifugation at 10,000 x g pelleted the agar, and the resulting supernatant was filter-sterilized using a Millex-HA 0.45 µm syringe-driven filter unit (Millipore, Billerica, MA), followed by storage at 4℃. Titers were obtained using serial dilutions of phage stock into SM, followed by the soft agar overlay technique described above and incubation at 30℃ until plaque formation was complete.

Host range analysis was performed using a panel of 27 clinical *S. maltophilia* and 19 *P. aeruginosa* strains. Soft-agar overlays containing 100 µl liquid culture were allowed to solidify

for 10 min at room temperature. Plates were spotted with 10 µl drops of DLP6 at multiple dilutions and assayed for clearing and/or plaque formation after incubation at 30℃ for 36 h.

For electron microscopy, phage stocks were prepared as described above with the following modifications: ½ LB agarose plates and ½ LB soft agarose were used for overlays, MilliQ-filtered water for phage recovery and a 0.22 µm filter was used for syringe-driven filtration. A carbon-coated copper grid was incubated with lysate for 2 min and stained with 4% uranyl acetate for 30 s. Transmission electron micrographs were captured using a Philips/FEI (Morgagni) transmission electron microscope with charge-coupled device camera at 80 kV (University of Alberta Department of Biological Sciences Advanced Microscopy Facility). The average capsid diameter, tail length and tail width were calculated using Microsoft Excel based on measurements from ten individual virions obtained using ImageJ.

Phage DNA isolation, RFLP analysis, and sequencing

DLP6 genomic DNA was isolated from bacteriophage lysate using the Wizard Lambda DNA purification system (Promega Corp., Madison, WI) with a modified protocol[168,265]. A 10 ml aliquot of high-titer filter-sterilized phage lysate was treated with 10 µl DNase I (Thermo Scientific, Waltham, MA), 100 µl 100x DNase I buffer (1 M Tris–HCl, 0.25 M $MgCl_2$, 10 mM $CaCl_2$), and 6 µl RNase (Thermo Scientific) and incubated 1 h at 37℃ to degrade the bacterial nucleic acids. Following incubation, 400 µl of 0.5 M EDTA and 25 µl of 20 mg/ml proteinase K (Applied Biosystems, Carlsbad, CA) was added and incubated 1 h at 55℃ to inactivate DNase I. The lysate was cooled to room temperature and added to 8.4 g of guanidine thiocyanate, along with 1 ml of 37℃ resuspended Wizard DNA Clean-Up Resin (Promega Corporation, Madison, WI). This mixture was rocked for 10 min then pelleted by centrifugation for 10 min at 5,000 x g. The supernatant was drawn off until ~5 ml remained. The remaining mixture was resuspended by swirling, transferred to a syringe attached to a Wizard Minicolumn (Promega Corporation). A Vac-Man Jr. Laboratory Vacuum Manifold (Promega Corporation) was used to filter supernatant under vacuum. The column was washed with 2 ml of 80% isopropanol and dried by centrifugation for 2 min at 10,000 x g. Phage DNA was eluted from the column following a 1 min incubation of 100 µl of 80℃ nuclease-free water (Integrated DNA Technologies, Coralville, IA) followed by centrifugation for 1 min at 10,000 x g. A NanoDrop ND-1000

spectrophotometer (Thermo Scientific, Waltham, MA) was used to determine the purity and concentration of eluted DNA.

Restriction fragment length polymorphism (RFLP) analysis was used with 19 FastDigest (Thermoscientific) restriction enzymes: HpaI, BamHI, EcoRI, AciI, HpaII, XbaI, HindIII, KpnI, SmaI, ApaI, SaII, PstI, SpHI, SacI, ClaI, NdeI, SpeI, XhoI, and HaeIII. Restriction reactions were set up using 1 µl of the FastDigest enzyme, 2 µl of the FastDigest restriction buffer, 1 µg of phage DNA and topped up to 20 µl with nuclease-free water. Reactions were incubated at 37℃ for 20 min and separated on a 1% (wt/vol) agarose gel in 1x TAE (pH 8.0). Sequencing of DLP6 was performed at The Applied Genomics Core at the University of Alberta. Purified DLP6 DNA was prepared for sequencing using a Nextera XT library prep kit, creating a library size of 223 bp. The library was used for paired-end sequencing on a MiSeq (Illumina, San Diego, CA) platform using a MiSeq v2 reagent kit. The Q30 for all reads was 92.5%.

Lifecycle determination of DLP6

DLP6 resistant colonies of *S. maltophilia* D1571 were isolated by a top agar overlay method using a phage stock at a titer of $1x10^5$ PFU/mL. Following overnight incubation at 30℃, a 3 mL aliquot of SM was transferred to each plate, and the supernatant was collected and used for serial dilutions to obtain superinfection resistant single colonies on ½ LB plates. Individual colonies were picked, washed with SM three times and used to produce freezer stocks. Superinfection experiments were performed using overnight cultures of the potential lysogens and DLP6 at a $1x10^5$ PFU/mL titer. Single colonies for each potential lysogen or pseudolysogen were used for colony PCR to detect the presence of DLP6. Identifying the temporary presence of DLP6 in the cell during pseudolysogeny was determined using specific sets of internal primers for DLP6.

Bioinformatics analysis

A single contig was assembled using the CLC Genomics Workbench (Qiagen, Toronto, ON). Although no ambiguous regions in the contig were observed, PCR amplification and sequencing were used to confirm the assembly. All attempts by PCR amplification to identify a DNA segment between the direct repeats or to show genome circularization via direct repeat annealing were negative. Open reading frames (ORFs) were identified with the GLIMMER plugin [172] for Geneious [217] using the Bacteria and Archaea setting, as well as GeneMarkS for

phage [174] and Prodigal [266]. Conserved domain searches were performed using CD-Search [175]. Pfam [224] was used to identify functions for hypothetical proteins from BLASTp. The contig was annotated and confirmed with an Interactive Remote Invocation Service utilizing the RAST pipeline [267-269]. BLASTn and BLASTp (for full genomes and individual proteins, respectively) were used to gain more information for each RAST annotation, and to identify any potentially related phages[176]. BLASTp results above $1.00E^{-3}$ were not recorded, and the coding sequence (CDS) was annotated as hypothetical. Rho-independent terminators were predicted using ARNold [270-272] searching both strands. Promoters were predicted using PHIRE [273] and plotted using WebLogo 3[274]. tRNAs were identified using tRNAscan-SE using the general tRNA model[222]. Multiple sequence alignments were performed with the top 216 BLASTp sequences for gp20 and from core T4 and cyanophage proteins [275] using the MUSCLE [180] plugin for Geneious. The maximum number of iterations selected was 8, with the anchor optimization option selected. The trees from iterations one and two were not retained. The distance measure for iteration one was kmer6_6 and was pctid_kimura for subsequent iterations. The clustering method was UPGMB for all iterations. An unrooted tree was constructed from MUSCLE alignments with the FastTree 2.1.5 [276,277] plugin for Geneious. The Jones-Taylor-Thornton model was used with rate category of sites set to 20. A PROmer comparison was conducted with DLP6 and FM12 with the following parameters: breaklen = 60, maxgap = 30, mincluster = 10, min-match = 3[178].

RESULTS AND DISCUSSION

Isolation, host range, and morphology

Phage DLP6 was isolated from a soil sample using clinical isolate *S. maltophilia* strain D1571 as the host. Propagation of DLP6 to high titer has proven difficult in liquid cultures, with liquid grown lysate concentrations remaining constant at 10^6 PFU/ml despite attempts to increase progeny numbers. DLP6 exhibits a unique plaquing inhibition that was previously observed in *Burkholderia cepacia* complex phages KL1 and AH2 [181], as well as *S. maltophilia* phage DLP1[163]. Although high titer stocks (10^{10} plaque forming units [PFU]/ml) can easily be obtained using the top agar plating method, use of such high titer stock inhibits the plaque formation on a bacterial lawn.

Plaquing of DLP6 is inhibited at titers above 10^6 plaque forming units (PFU/ml). Plaque development occurs readily at 30℃ within 24 h, forming diffuse plaques with irregular borders and a mean size of 0.8 ± 0.3 mm. Host range analysis of DLP6 revealed a moderate host range within *S. maltophilia* clinical isolates, infecting 13 out of 27 clinical isolates (Table 1-2). Whereas *S. maltophilia* phages DLP1 and DLP2 exhibited some cross-species infectivity 22, extended host range analysis of DLP6 using *P. aeruginosa* isolates did not yield successful infections. Initially, we produced evidence to suggest that DLP6 existed in *S. maltophilia* D1571 as a prophage. However, after extensive experimentation, it was determined that DLP6 undergoes pseudolysogeny. PFGE analysis using SpeI or XbaI separately showed no integration of the DLP6 genome into the *S. maltophilia* D1571 chromosome, and DLP6-specific PCR indicated the genome's presence after 2–3 passages, but not after >5 passages. DLP6 is classified in the *Myoviridae* family of the *Caudovirales* order due to its icosahedral head and contractile tail (Figure 6-1). The average capsid height, tail length and width measurements for DLP6 are 99, 144, and 23 nm respectively.

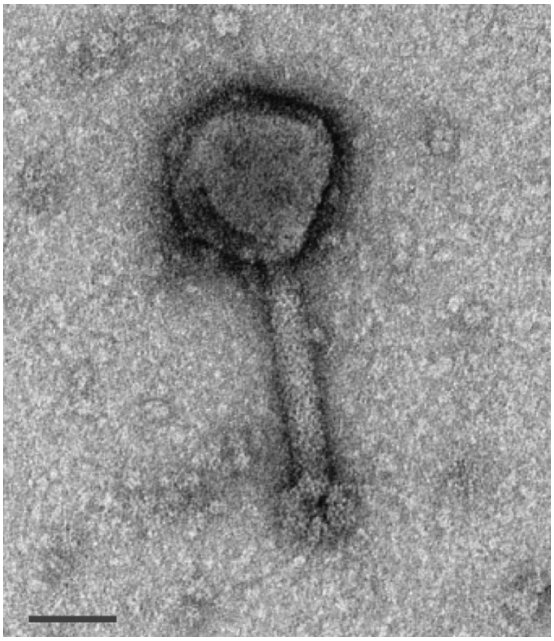

Figure 6-1: DLP6 phage morphology. Liquid phage lysate was incubated on a carbon coated copper grid, stained with 4% uranyl acetate and visualized at 180,000-fold magnification by a transmission electron microscope. Scale bar represent 50 nm. The average capsid height measurement for DLP6 was 99 nm, average tail length of 144 nm and average tail width of 23 nm.

Table 6-1: Extended host range analysis of DLP6

S. maltophilia strain	DLP6	*P. aeruginosa* strain	DLP6
101[c]	-	PA01	–
102[c]	+	HER1004	–
103[c]	+	HER1012	–
152[c]	-	14715	–
155[c]	+++	Utah3	–
174[c]	-	Utah4	–
176[c]	+	14655	–
213[c]	+++	6106	–
214[c]	-	pSHU-OTE	–
217[c]	++	D1606D[a,b]	–
218[c]	-	D1615C[a,b]	–
219[c]	+	D1619M[a,b]	–
230[c]	+	D1620E[a,b]	–
236[c]	-	D1623C[a,b]	–
242[c]	-	ENV003[a]	–
249[c]	-	ENV009[a]	–
278[c]	-	FC0507[a]	–
280[c]	-	R285	–
282[c]	-	14672	–
287[c]	++		
446[c]	+		
667[c]	++		
D1585[a,b]	-		
D1571[a,b]	+++		
D1614[a,b]	-		
D1576[a,b]	++		
D1568[a,b]	-		

–, No sensitivity to phage; +, plaques at 10-2; ++, clearing at 10-2; +++, plaques at 10-4; ++++, plaques at 10-6.
a Obtained from the Canadian *Burkholderia cepacia* complex Research Referral Repository.
b Cystic fibrosis patient isolate.
c Isolates from the Provincial Laboratory for Public Health - North (Microbiology), Alberta Health Services.

Genome characterization of DLP6

Purified DLP6 gDNA was isolated and exposed to a panel of 19 restriction enzymes for RFLP analysis, but only HpaII was capable of digesting DLP6 gDNA. The DLP6 genome was assembled into a linear scaffold of 168,489 bp with a GC content of 55.8% using 43,112 reads

for a mean coverage of 57 reads and an overall Q30 score of 93.1%. The DLP6 genome can be found in GenBank with the accession number KU682439.2. The genome is predicted to encode 241 open reading frames and 30 tRNAs of 14 different specificities (Figure 6-2, Table 6-2 and Table 6-3). The genome is arranged in a semi-modular format, with genes encoding DNA replication and repair proteins (dark blue) primarily grouped, whereas the regulatory proteins (medium blue) are grouped within the same region as the DNA replication and repair proteins. Auxiliary metabolic genes (yellow) are dispersed throughout the genome, occurring in small pairs rather than a large set. The phage morphogenesis genes (orange) are grouped in a cluster, except for the encoded gp34 (AAY80_073; long tail fiber). The 30 tRNA coding regions (pink) are grouped spanning the genome from 92,962–113,834 bp. There is no lysis (light blue) module in DLP6; instead, four genes encoding typical lysis proteins are randomly located throughout the genome.

Three interesting phage-encoded proteins are ADP-ribosyltransferases (Alt: AAY80_209, ModA: AAY80_029, ADP-ribosyltransferase: AAY80_145) that were identified using the Pfam database. Two of these proteins are orthologs to T4 proteins Alt and ModA. It is known that phage T4 encodes three ADP-ribosyltransferases (Alt, ModA, and ModB), each modifying specific groups of host proteins. The Alt protein is a component of the phage head and enters the host cell during the infection process with the phage DNA, ModA and ModB[278]. Following entry into the cell, Alt immediately ADP-ribosylates the host RNA polymerase (RNAP), causing transcription of host genes to stop and transcription from the T4 “early” promoters to be carried out instead[279]. The ModA modification of the host RNAP prevents transcription from T4 early promoters and possibly primes the RNAP for T4-encoded auxiliary factors (gp55, gp33, gp45, and gp44/62) to transcribe middle and late genes[280]. Identification of the previously hypothetical proteins into the Alt and ModA families helps to provide insight into the possible role these proteins play in DLP6 phage infection and transcription initiation and regulation.

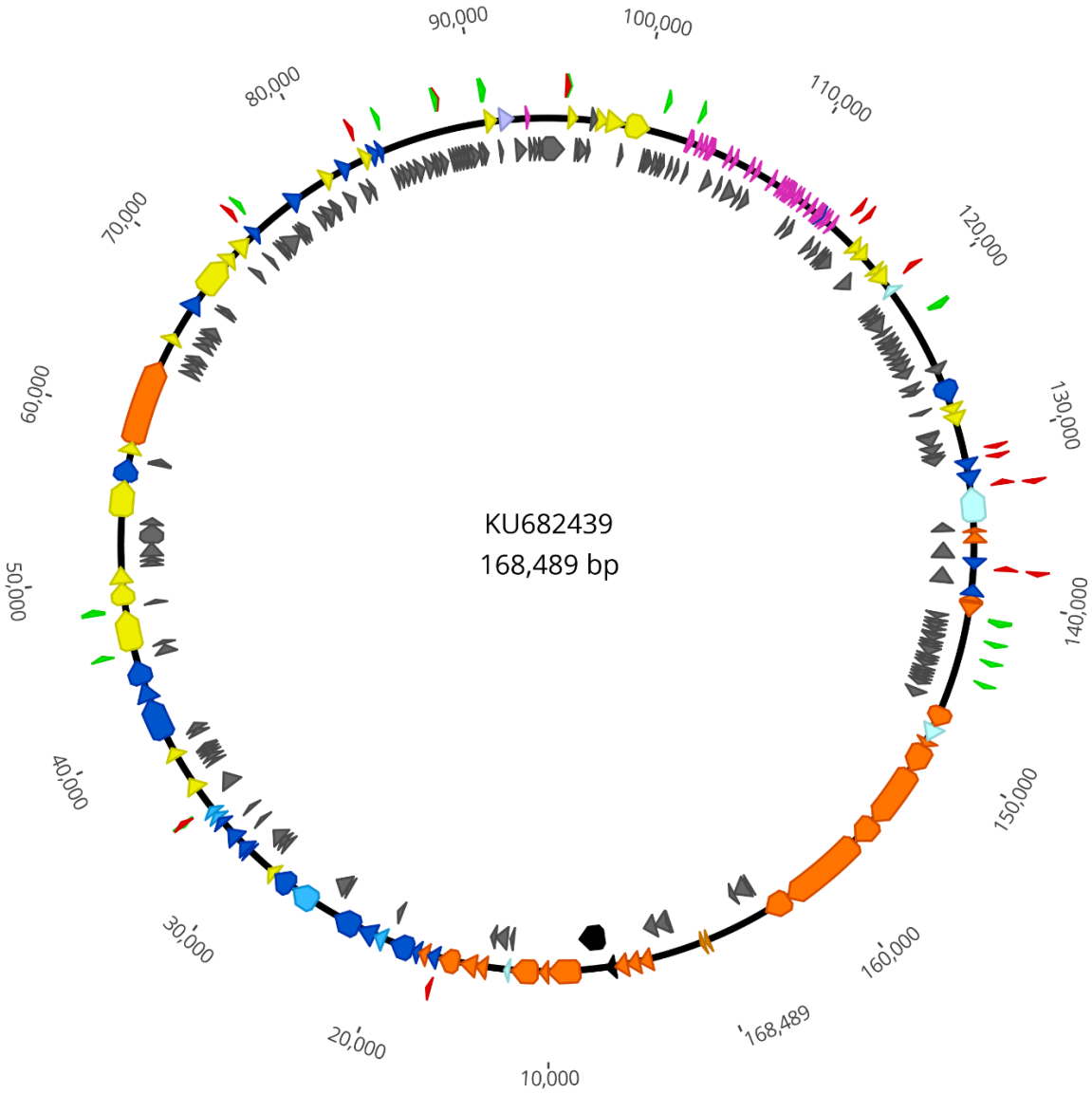

Figure 6-2: Genome map of DLP6. The scale in bp is shown on the outermost periphery of the map with the late viral promoters (green) and Rho-independent terminators (red). Assigned functions for each predicted open reading frame are as follows: auxiliary metabolic genes; yellow, DNA packaging; black, DNA replication & repair; dark blue, regulatory; medium blue, hypothetical; grey, lysis; light blue, repeat region; pale orange, transposase; lilac, tRNA; pink, and virion morphogenesis; orange. Due to space constraints, most hypothetical genes are located in the inner periphery of the circle to reduce overlap.

Phage promoter sequences lack the conserved structure observed in bacterial promoters with -10 and -35 regions; instead, they feature short consensus sequences that are specific to different phages[281]. These short consensus sequences were identified using the phage-specific program PHIRE and visualized with WebLogo 3 (Figure 6-3). There are 27 phage promoters identified, with 24 of the phage promoters found repeating in groups of two in front of a gene cluster. A single phage promoter is located upstream of genes AAY80_058 (hypothetical protein) and AAY80_059 (peptidase protein). The next single phage promoter is located upstream of the gene cluster beginning with AAY80_140 (kinase protein) though to AAY80_151 (hypothetical protein). The last single phage promoter is located upstream of a small cluster of hypothetical proteins encoded from AAY80_217 though to AAY80_220. Two phage promoters are found upstream of AAY80_060 (hypothetical protein) in a gene cluster coding for 32 proteins, including DNA primase (AAY80_083) phage tail fiber (AAY80_073), ribonucleotide diphosphate reductase subunit alpha (AAY80_086) and beta (AAY80_089), and many hypothetical proteins (locus tags ending in 060, 063–067, 070–071, 074–077, 079–082, 084–085, 088 and 090).

The next set of phage promoters is located upstream of AAY80_092 (hypothetical protein) to AAY80_112 (hypothetical protein). Annotated genes included in this gene cluster are AAY80_098 (RNase H), AAY80_104 (PhoH), AAY80_106 (exonuclease), AAY80_109 (SleB), AAY80_110 (dCMP deaminase) and AAY80_111 (deoxycytidylate deaminase). The next two sets of gene clusters, AAY80_113–120 and AAY80_121–130, encode hypothetical proteins only and both clusters are under the control of two promoters each. Gene cluster AAY80_131–138 utilizes two promoter sequences and encodes mainly hypothetical proteins, but also a guanosine 3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (AAY80_131) and a transposase (AAY80_133). The next gene cluster under control of two promoters is upstream of AAY80_151 (hypothetical protein) though to AAY80_158 (hypothetical protein) and contains nine tRNAs. The remaining 21 tRNAs are under control of two promoters upstream of AAY80_159 (hypothetical protein) through to AAY80_180 (hypothetical protein). Two promoters control the next gene cluster spanning from AAY80_181–202 that contains genes encoding many hypothetical proteins and proteins such as DNA ligase (AAY80_192), DNA helicase loader (AAY80_201) and ssDNA binding protein (AAY80_202). Three small gene clusters encoding a total of 12 hypothetical proteins (AAY80_213–216, AAY80_217-220, and AAY80_221–224) are under control of two

phage promoters. The last set of double phage promoters controls the gene cluster from AAY80_225 to AAY80_057 where the phage genome is circularized. This final gene cluster encodes many proteins involved in DNA replication and homologous recombination.

Table 6-2: DLP6 tRNA predictions.

tRNA	Strand	Start	End	Type	Anticodon	Cove Score
1	+	92,962	93,035	Arg	CCT	61.34
2	+	103,314	103,387	Pro	TGG	73.92
3	+	103,398	103,469	Ile	GAT	43.6
4	+	103,950	104,034	Leu	TAG	55.46
5	+	104,297	104,368	Gly	GCC	74.95
6	+	104,656	104,729	Leu	CAA	50.06
7	+	104,738	104,829	Ile	TAT	46.18
8	+	104,845	104,917	Phe	GAA	44.78
9	+	106,026	106,099	Arg	CCG	64.25
10	+	106,481	106,557	Th	CGT	68.07
11	+	107,530	107,612	Pseudo	GTT	29
12	+	107,978	108,050	Lys	TTT	55.23
13	+	109,099	109,169	Ile	GAT	63.46
14	+	109,728	109,799	Lys	CTT	56.9
15	+	110,017	110,089	Ala	TGC	68.8
16	+	110,184	110,255	Cys	GCA	47.48
17	+	110,348	110,418	Gln	CTG	63.37
18	+	110,432	110,504	Trp	CCA	66.76
19	+	110,516	110,587	Val	TAC	39.72
20	+	110,832	110,903	Th	GGT	74.72
21	+	111,078	111,149	Val	GAC	67.9
22	+	111,579	111,652	Arg	TCT	56.91
23	+	112,066	112,139	Leu	GAG	56.11
24	+	112,149	112,221	Phe	GAA	74.14
25	+	112,574	112,645	His	GTG	44.49
26	+	112,656	112,728	Lys	CTT	75.8
27	+	112,740	112,813	Arg	ACG	69.71
28	+	113,092	113,166	Pro	CGG	65.21
29	+	113,330	113,403	Leu	CAG	56.95
30	+	113,763	113,834	Sup	CTA	48.3

Table 6-3: DLP6 genome annotation. BLASTp results above 1.00E-3 were not recorded and the coding sequence was annotated as hypothetical. Coding sequences without BLASTp results were left blank for closest relative, E-value, source, and accession.

CDS	Coding region	Length (AA)	Start Codon	Putative Function	Closest relative	E-Value	Source	Accession
1	1664 - 1837	57	ATG	hypothetical protein				
2	1821 - 2699	292	ATG	hypothetical protein	hypothetical protein	2.00E-05	Φpp2	WP_018251053.1
3	2696 - 3448	250	ATG	hypothetical protein				
4	3483 - 4223	246	ATG	gp13: neck protein	neck protein	3.00E-101	Cr30	YP_009098960.1
5	4227 - 4982	251	ATG	gp14: neck protein	neck protein	3.00E-62	ΦN3	YP_009212340.1
6	4982 - 5785	267	ATG	gp15: proximal tail sheath stabilization	proximal tail sheath stabilization	2.00E-68	ΦM12	YP_009143012.1
7	5782 - 6300	172	ATG	g16: small terminase subunit	small terminase subunit	4.00E-22	S-SM1	YP_004323004.1
8	6272 - 7936	554	ATG	gp17: large terminase subunit	large terminase subunit	0.00E+00	ACG-2014c	YP_007001827.1
9	7999 - 10014	671	ATG	gp18: tail sheath protein	tail sheath protein	0.00E+00	Cr30	YP_009098945.1
10	10035 - 10568	177	ATG	gp19: tail tube protein	tail tube protein	1.00E-51	ΦM12	YP_009142981.1
11	10636 - 12303	555	ATG	gp20: portal vertex of head	portal vertex of head protein	0.00E+00	S-SM1	YP_004323020.1
12	12329 - 12490	53	ATG	hypothetical protein				
13	12490 - 12747	85	ATG	holin protein	holin	6.00E-03	JCL1032	YP_007002991.1

14	12750 - 13535	261	ATG	hypothetical protein				
15	13538 - 13855	105	ATG	hypothetical protein				
16	13855 - 14517	220	ATG	gp21: prohead core scaffold and protease	prohead core scaffold and protease	3.00E-75	S-PM2	YP_195140.1
17	14565 - 15578	337	ATG	gp22: scaffold prohead core	scaffold prohead core	9.00E-55	uvMED	BAR30798.1
18	15618 - 16919	433	ATG	gp23: precursor of major head subunit	precursor of major head subunit	0.00E+00	ΦN3	YP_009212304.1
19	17030 - 17623	197	GTG	DexA	exonuclease	3.00E-13	44RR2.8t	NP_932376.1
20	17660 - 18169	169	ATG	gp3: tail tube protein	phage tail tube protein	2.00E-32	uvMED	BAR32912.1
21	18185 - 18619	144	ATG	UvsY	putative repair and recombination protein	4.00E-33	ΦM12	YP_009142965.1
22	18623 - 20086	487	GTG	UvsW	UvsW	2.00E-157	uvMED	BAR36206.1
23	20083 - 20469	128	GTG	hypothetical protein				
24	20579 - 21211	210	ATG	gp55: sigma factor late transcription	hypothetical protein	3.00E-48	P-RSM6	YP_007675119.1
25	21247 - 22335	362	ATG	gp47: recombination endonuclease subunit	recombination endonuclease subunit	3.00E-106	ΦN3	YP_009212290.1
26	22332 - 24047	571	GTG	gp46: recombination endonuclease	recombination endonuclease subunit	6.00E-176	S-ShM2	YP_004322802.1

27	24044 - 24238	64	ATG	hypothetical protein	polyprotein	8.00E-04	ATROP06/BR/2012	AKI82132.1
28	24235 - 25098	287	ATG	hypothetical protein				
29	25452 - 27143	563	ATG	ModA	hypothetical protein	8.00E-19	RSL1	YP_001950115.1
30	27179 - 28525	448	ATG	SNF2 DNA repair protein	SNF2 DNA repair protein	4.00E-122	Cr30	YP_009098887.1
31	28522 - 29079	185	ATG	peptide deformylase	peptide deformylase	4.00E-43	Cr30	YP_009098904.1
32	29079 - 29429	116	ATG	hypothetical protein	hypothetical protein	1.00E-13	*Gemmata sp.* IIL30	WP_052559698.1
33	29429 - 29731	100	ATG	hypothetical protein				
34	29740 - 30597	285	ATG	hypothetical protein	hypothetical protein	1.00E-121	Cr30	YP_009098901.1
35	30619 - 30861	80	ATG	DsbA dsDNA binding protein	DsbA dsDNA binding protein	3.00E-05	T4	NP_049858.1
36	30877 - 31548	223	ATG	gp45: sliding clamp	sliding clamp	3.00E-67	P-SSM2	YP_214389.1
37	31598 - 31771	57	ATG	hypothetical protein				
38	31764 - 32699	311	ATG	gp44: sliding clamp loader subunit	clamp loader subunit	7.00E-130	S-MbCM25	AHB80773.1
39	32696 - 32929	77	ATG	hypothetical protein				
40	33166 - 33600	144	GTG	gp62: clamp loader subunit	clamp loader subunit	3.00E-33	ΦM12	YP_009142948.1
41	33597 - 33989	130	ATG	regA	regA	2.00E-54	S-IOM18	YP_008126450.1
42	33989 - 34405	138	ATG	MazG	triphosphate pyrophosphohydrolase	4.00E-23	*M. goulette*	AGF85421.1

43	34789 - 35745	318	ATG	hypothetical protein	hypothetical protein	2.00E-82	S13	YP_009196385.1
44	35745 - 36674	309	ATG	N-acetyltransferase	N-acetyltransferase	2.00E-08	*A. inops*	WP_052131127.1
45	36667 - 36897	76	ATG	hypothetical protein				
46	37045 - 37290	81	ATG	hypothetical protein	hypothetical protein	8.00E-07	Lu11	YP_006382802.1
47	37301 - 37525	74	ATG	hypothetical protein				
48	37522 - 37737	71	ATG	hypothetical protein	hypothetical protein	1.00E-07	uvMED	BAR33778.1
49	37737 - 37886	49	ATG	hypothetical protein				
50	37890 - 38150	86	ATG	hypothetical protein				
51	38223 - 38837	204	ATG	sulfotransferase	sulfotransferase	2.00E-28	*Lysobacter*	WP_055903507.1
52	38845 - 39264	139	ATG	hypothetical protein				
53	39261 - 39680	139	ATG	hypothetical protein	hypothetical protein	4.00E-38	*J. lividum*	WP_010396923.1
54	39682 - 42228	848	ATG	gp43: DNA polymerase	DNA polymerase	0.00E+00	syn9	YP_717843.1
55	42267 - 43361	364	ATG	UvsX	recombination protein	3.00E-173	IME-SM1	AKO61686.1
56	43361 - 44788	475	ATG	gp41: DNA primase/helicase	DNA primase/helicase	0.00E+00	S-IOM18	YP_008126466.1
57	44785 - 45309	174	ATG	hypothetical protein	hypothetical protein	2.00E-06	Cr30	YP_009098845.1
58	45427 - 45597	56	ATG	hypothetical protein				

59	45607 - 48036	809	ATG	peptidase	peptidase	2.00E-75	P-RSM4	YP_004323283.1
60	48172 - 48441	89	ATG	hypothetical protein				
61	48452 - 49732	426	ATG	CobS	porphyrin biosynthetic protein	9.00E-119	ACG-2014i	YP_009140910.1
62	49818 - 50792	324	ATG	Td	FAD-dependent thymidylate synthase	8.00E-120	VCM	CUR44270.1
63	50789 - 51145	118	GTG	hypothetical protein	hypothetical protein	1.00E-18	7-7-1	YP_007006492.1
64	51138 - 51320	60	ATG	hypothetical protein				
65	51394 - 52311	305	ATG	hypothetical protein	hypothetical protein	3.00E-68	ΦM12	YP_009143274.1
66	52316 - 53551	411	ATG	hypothetical protein	hypothetical protein	2.00E-87	HTVC008M	YP_007517993.1
67	53580 - 53993	137	ATG	hypothetical protein				
68	54039 - 56207	722	ATG	cytitidyltransferase	cytitidyltransferase	9.00E-37	ACG-2014a	AIX45294.1
69	56234 - 57535	433	ATG	gp49: exodeoxyribonuclease VII large subunit	exodeoxyribonuclease VII large subunit	2.00E-03	*T. aerophilum*	WP_006460299.1
70	57535 - 57687	50	ATG	hypothetical protein				
71	57687 - 57998	103	ATG	hypothetical protein	hypothetical protein	3.00E-05	uvMED	BAR32939.1
72	58009 - 58653	214	ATG	2OG-Fe(II) oxygenase	2OG-Fe(II) oxygenase	2.00E-20	*Phycodnavirus* 1	YP_009174691.1
73	58653 - 63980	1775	ATG	gp34: phage tail fiber	hypothetical protein	1.00E-43	11b	YP_112522.1

74	63977 - 64384	135	ATG	hypothetical protein	hypothetical protein	2.00E-03	*S. panacihumi*	WP_057645976.1
75	64388 - 64858	156	ATG	hypothetical protein	hypothetical protein	2.00E-14	*S. maltophilia*	WP_049447215.1
76	64891 - 65058	55	ATG	hypothetical protein				
77	65068 - 65448	126	ATG	hypothetical protein	hypothetical protein	1.00E-32	ΦM12	YP_009143268.1
78	65450 - 65956	168	ATG	2OG-Fe(II) oxygenase	2OG-Fe(II) oxygenase	2.00E-19	S-RSM4	YP_003097374.1
79	65960 - 66484	174	ATG	hypothetical protein	hypothetical protein	2.00E-30	ΦM12	YP_009143267.1
80	66481 - 66849	122	ATG	hypothetical protein	hypothetical protein	3.00E-09	1M3-16	YP_009037295.1
81	66846 - 67148	100	GTG	hypothetical protein				
82	67145 - 67717	190	ATG	hypothetical protein	hypothetical protein	1.00E-14	HTVC008M	YP_007518013.1
83	67787 - 68827	346	ATG	gp61: DNA primase	DNA primase	2.00E-104	ΦM12	YP_009143264.1
84	68824 - 69006	60	ATG	hypothetical protein	hypothetical protein	6.00E-06	ACG-2014h	YP_009008190.1
85	68999 - 69253	84	GTG	hypothetical protein				
86	69231 - 71531	766	ATG	nrdA	ribonucleotide diphosphate reductase subunit alpha	0.00E+00	vB_CsaM_GAP32	YP_006987403.1
87	71543 - 72187	214	ATG	methyltransferase	methyltransferase	4.00E-42	ΦN3	YP_009212249.1
88	72206 - 72553	115	ATG	hypothetical protein	hypothetical protein	5.00E-07	ΦCP26F	YP_007004033.1

89	72553 - 73599	348	ATG	nrdB	ribonucleotide diphosphate reductase subunit beta	2.00E-128	vB_CsaM_GAP32	YP_006987405.1
90	73596 - 73829	77	ATG	hypothetical protein				
91	73893 - 74438	181	ATG	RuvC	hypothetical protein	1.00E-36	Cr30	YP_009098799.1
92	74552 - 74779	75	ATG	hypothetical protein				
93	74803 - 75054	83	ATG	hypothetical protein				
94	75054 - 76184	376	ATG	hypothetical protein	hypothetical protein	5.00E-31	Cr30	YP_009098763.1
95	76184 - 76414	76	ATG	hypothetical protein				
96	76416 - 76646	76	ATG	hypothetical protein				
97	76679 - 76990	103	ATG	hypothetical protein				
98	76987 - 77865	292	GTG	RNase H	Rnase H	5.00E-83	S-CBM2	AFK66374.1
99	77876 - 78283	135	GTG	T4 Gc 313: hypothetical protein	hypothetical protein	3.00E-21	Cr30	YP_009099041.1
100	78293 - 78475	60	ATG	hypothetical protein				
101	78488 - 79075	195	ATG	hypothetical protein	hypothetical protein	4.00E-34	ΦN3	YP_009212443.1
102	79072 - 79266	64	ATG	hypothetical protein				
103	79273 - 79512	79	ATG	T4 Gc 321: hypothetical protein	hypothetical protein	9.00E-17	ΦM12	YP_009143120.1

104	79532 - 80239	235	ATG	PhoH	PhoH	5.00E-70	P-SSM3	YP_008130050.1
105	80257 - 80751	164	ATG	hypothetical protein	hypothetical protein	6.00E-40	CcrColossus	YP_006988422.1
106	80732 - 81448	238	GTG	exonuclease	hypothetical protein	7.00E-43	S-CBM2	AFK66381.1
107	81475 - 81945	156	ATG	hypothetical protein	hypothetical protein	1.00E-09	Cr30	YP_009098845.1
108	82033 - 82272	79	ATG	hypothetical protein	hypothetical protein	1.00E-06	uvMED	BAR35113.1
109	82346 - 82906	186	ATG	SleB	SleB	1.00E-22	ΦTE	YP_007392495.1
110	82910 - 83398	162	ATG	dCMP deaminase	dCMP deaminase	5.00E-39	vB_PaeS_PA01_Ab18	YP_009125122.1
111	83403 - 83759	118	ATG	deoxycytidylate deaminase	deoxycytidylate deaminase	2.00E-10	TMA	YP_004782245.1
112	83743 - 84003	86	ATG	gp33: hypothetical protein	hypothetical protein	1.00E-06	Cr30	YP_009099022.1
113	84160 - 84465	101	ATG	hypothetical protein				
114	84467 - 84892	141	ATG	hypothetical protein				
115	84892 - 85305	137	ATG	hypothetical protein				
116	85382 - 85930	182	ATG	hypothetical protein				
117	85955 - 86509	184	ATG	hypothetical protein				
118	86519 - 86731	70	ATG	hypothetical protein	hypothetical protein	7.00E-13	*Oscillibacter sp.* ER4	WP_008981440.1
119	86737 - 87144	135	ATG	hypothetical protein	hypothetical protein	1.00E-23	vB_CsaM_GAP31	YP_006986953.1

120	87155 - 87586	143	ATG	hypothetical protein				
121	87732 - 88007	91	ATG	hypothetical protein				
122	88009 - 88188	59	ATG	hypothetical protein				
123	88193 - 88435	80	ATG	hypothetical protein				
124	88432 - 88635	67	ATG	hypothetical protein				
125	88632 - 88772	46	ATG	hypothetical protein				
126	88769 - 89026	85	GTG	hypothetical protein	hypothetical protein	3.00E-18	Jay2Jay	AIW02656.1
127	89023 - 89274	83	ATG	hypothetical protein	hypothetical protein	6.00E-07	UFV-P2	YP_007518481.1
128	89264 - 89725	153	ATG	hypothetical protein				
129	89725 - 90057	110	ATG	hypothetical protein				
130	90050 - 90364	104	ATG	hypothetical protein				
131	90484 - 91092	202	ATG	(p)ppGpp synthase II	hypothetical protein	1.00E-34	*P. stutzeri*	WP_052813543.1
132	91092 - 91295	67	ATG	hypothetical protein				
133	91298 - 92203	301	ATG	transposase	transposase	1.00E-71	*P. mendocina*	WP_047586214.1
134	92271 - 92855	194	ATG	hypothetical protein	hypothetical protein	6.00E-16	Φ3	YP_009207497.1
135	93056 - 93361	101	ATG	hypothetical protein				

136	93467 - 93736	89	ATG	hypothetical protein				
137	93746 - 93976	76	ATG	hypothetical protein				
138	94038 - 95438	466	ATG	hypothetical protein	hypothetical protein	2.00E-161	S13	YP_009196578.1
139	95658 - 96224	188	ATG	Tk: kinase protein	kinase protein	5.00E-08	Cd1	ADD21639.1
140	96234 - 96440	68	ATG	hypothetical protein	hypothetical protein	6.00E-13	Cr30	YP_009098992.1
141	96443 - 97003	186	ATG	hypothetical protein	hypothetical protein	2.00E-23	Φpto-bp6g	YP_009015268.1
142	96993 - 97181	62	ATG	hypothetical protein	hypothetical protein	8.00E-17	SFP10	YP_004895215.1
143	97159 - 97410	83	ATG	hypothetical protein	hypothetical protein	9.00E-05	*Microbacterium sp.* No. 7	WP_054686480.1
144	97403 - 98062	219	ATG	metallophoesterase	metallophoesterase	1.00E-18	*M. denitrificans*	WP_036280110.1
145	98136 - 99215	359	ATG	ADP-ribose pyrophosphatase	ADP-ribose pyrophosphatase	6.00E-112	IME-SM1	AKO61723.1
146	99215 - 99215	32	ATG	hypothetical protein				
147	99327 - 100772	481	ATG	nicotinamide phosphoribosyl-transferase	nicotinamide phosphoribosyl-transferase	4.00E-149	IME-SM1	AKO61724.1
148	100772 - 100963	63	ATG	hypothetical protein				
149	100960 - 101154	64	GTG	hypothetical protein				
150	101151 - 101396	81	ATG	hypothetical protein	hypothetical protein	2.00E-13	nt-1	YP_008125398.1
151	101599 - 101889	96	GTG	hypothetical protein				

152	101906 - 102097	63	ATG	hypothetical protein				
153	102104 - 102313	69	ATG	hypothetical protein	hypothetical protein	2.00E-05	*Alcanivorax sp.* 43B_GOM-46m	WP_026949135.1
154	102610 - 102834	74	ATG	hypothetical protein				
155	103134 - 103304	56	ATG	hypothetical protein				
156	103728 - 103940	70	GTG	hypothetical protein	hypothetical protein	2.00E-05	*N. caesariensis*	WP_007019913.1
157	105246 - 105767	173	ATG	hypothetical protein				
158	106212 - 106472	86	ATG	hypothetical protein				
159	106759 - 107520	253	ATG	hypothetical protein	hypothetical protein	9.00E-14	*Pseudomonas sp*. Leaf58	WP_056798772.1
160	107664 - 107888	74	ATG	hypothetical protein				
161	108080 - 108562	160	GTG	hypothetical protein	hypothetical protein	5.00E-62	*Candidatus Methylopumilus planktonicus*	WP_046488394.1
162	111363 - 111572	69	ATG	hypothetical protein				
163	111791 - 112057	88	ATG	hypothetical protein				
164	112831 - 113079	82	ATG	GroES	GroES molecular chaperone protein	4.00E-03	*B. retamae*	WP_057842828.1
165	113397 - 113756	119	ATG	hypothetical protein	hypothetical protein	1.00E-25	*F. filum*	WP_035655512.1

166	114007 - 114354	115	ATG	hypothetical protein				
167	114587 - 114757	56	ATG	hypothetical protein	hypothetical protein	1.00E-05	*Nitrosomonas sp.* AL212	ADZ27808.1
168	114750 - 114974	74	ATG	hypothetical protein				
169	115030 - 115647	205	ATG	hypothetical protein	hypothetical protein	2.00E-07	*H. adhaerens*	WP_035570500.1
170	115647 - 116264	205	ATG	methyltransferase	methyltransferase	2.00E-12	uvMED	BAQ89138.1
171	116266 - 116916	216	ATG	nucleotide-diphospho-sugar transferase	nucleotide-diphospho-sugar transferase	7.00E-32	CcrColossus	YP_006988363.1
172	116909 - 117700	263	GTG	sialyltransferase				
173	117697 - 117957	86	ATG	nrdC	glutaredoxin	3.00E-17	Cr30	YP_009098769.1
174	117973 - 118926	317	ATG	glycosyl transferase protein	glycosyl transferase protein	3.00E-81	ΦM12	YP_009143113.1
175	119098 – 119577[a]	159	ATG	endolysin	endolysin	2.00E-46	RL-2015	AJG41873.1
176	119690 - 119914	74	ATG	hypothetical protein				
177	119901 - 120218	105	ATG	hypothetical protein				
178	120220 - 120411	63	ATG	hypothetical protein				
179	120428 - 121342	304	ATG	hypothetical protein				
180	121339 - 121599	86	ATG	hypothetical protein				

181	121741 - 122064	107	ATG	hypothetical protein				
182	122068 - 122268	66	ATG	hypothetical protein				
183	122270 - 122596	108	ATG	hypothetical protein	hypothetical protein	7.00E-08	P12053L	YP_006560897.1
184	122596 - 122985	129	ATG	hypothetical protein	hypothetical protein	6.00E-18	*C. botulinum*	WP_024931995.1
185	123031 - 123417	128	ATG	hypothetical protein				
186	123417 - 123779	120	ATG	hypothetical protein	hypothetical protein	1.00E-27	*Chyseobacterium sp.* BLS98	KMQ60445.1
187	123786 - 124307	173	ATG	hypothetical protein				
188	124307 - 124864	185	ATG	hypothetical protein	hypothetical protein	2.00E-06	P35	YP_001468837.1
189	124839 - 125438	199	ATG	hypothetical protein	hypothetical protein	3.00E-09	P35	YP_001468837.1
190	125460 - 125837	125	ATG	hypothetical protein	hypothetical protein	3.00E-26	ΦM12	YP_009143283.1
191	125834 - 125983	49	GTG	hypothetical protein				
192	125986 - 127338	450	ATG	gp30: DNA ligase	hypothetical protein	9.00E-117	*Methylibium petroleiphilum*	WP_011831629.1
193	127349 - 127516	55	ATG	hypothetical protein				
194	127516 - 128067	183	ATG	Fe(II)-dependent oxygenase	Fe(II)-dependent oxygenase	1.00E-16	uvMED	BAQ86661.1
195	128070 - 128867	265	ATG	Ser/Th phosphatase protein	Ser/Th phosphatase protein	5.00E-43	PBECO_4	YP_009150422.1

196	128864 - 129058	64	ATG	hypothetical protein				
197	129043 - 129774	243	ATG	hypothetical protein	hypothetical protein	6.00E-69	Sano	AHB12055.1
198	129771 - 130301	176	GTG	hypothetical protein				
199	130330 - 130521	63	ATG	hypothetical protein				
200	130567 - 131070	167	ATG	hypothetical protein	hypothetical protein	6.00E-39	Cr30	YP_009098835.1
201	131151 - 131765	204	ATG	gp59: DNA helicase loader	DNA helicase loader	1.00E-40	ΦN3	YP_009212381.1
202	131859 - 132815	318	ATG	gp32: ssDNA binding protein	ssDNA binding protein	9.00E-108	P-SSM2	ACY75884.1
203	132883 – 135045[a]	720	ATG	lysozyme	lysozyme	3.00E-29	uvMED	BAR34129.1
204	135049 – 135495[a]	148	ATG	hypothetical protein				
205	135482 – 135664[a]	60	ATG	gp51: baseplate hub assembly catalyst protein	hypothetical protein	4.00E-15	S-SKS1	YP_007674616.1
206	135667 – 136368[a]	233	ATG	gp26: baseplate hub subunit protein	hypothetical protein	3.00E-44	Cr30	YP_009099008.1
207	136377 – 137273[a]	298	ATG	19.2: hypothetical protein	hypothetical protein	3.00E-26	Cr30	YP_009099003.1
208	137378 – 137950	190	ATG	endonuclease protein	endonuclease protein	3.00E-04	WS6 bacterium 34_10	KUK77195.1
209	137986 – 138999[a]	337	GTG	Alt	hypothetical protein	4.00E-05	*Clostridium sp.* CAG:813	CDF00004.1
210	138996 – 139613[a]	205	ATG	gp2: DNA end protector protein	DNA end protector protein	3.00E-65	ΦM12	YP_009143037.1

211	139594 – 140040[a]	148	ATG	gp4: head completion protein	head completion protein	1.00E-51	P-SSM2	YP_214244.1
212	140085 – 140987	300	ATG	gp48: baseplate tail tube cap protein	baseplate tail tube cap protein	8.00E-18	ΦEa2809	YP_009147536.1
213	141136 – 141477	113	ATG	hypothetical protein	hypothetical protein	6.00E-06	*Yersinia pekkaneii*	WP_049615257.1
214	141495 – 141749	84	ATG	hypothetical protein				
215	141755 – 142057	100	ATG	hypothetical protein				
216	142054 – 142308	84	ATG	hypothetical protein				
217	142418 – 142588	56	ATG	hypothetical protein				
218	142588 – 142962	124	ATG	hypothetical protein				
219	142959 – 143174	71	ATG	hypothetical protein				
220	143179 – 143391	70	ATG	hypothetical protein				
221	143486 – 144034	182	ATG	hypothetical protein	hypothetical protein	4.00E-18	*S. maltophilia*	WP_053451543.1
222	144036 – 144227	63	ATG	hypothetical protein	hypothetical protein	3.00E-06	ΦL7	YP_002922638.1
223	144238 – 144465	75	ATG	hypothetical protein				
224	144462 – 144728	88	ATG	hypothetical protein				
225	144850 – 145041	63	ATG	hypothetical protein				
226	145055 – 145288	77	ATG	hypothetical protein				

227	145285 – 145521	78	GTG	hypothetical protein				
228	145514 – 145717	67	ATG	hypothetical protein				
229	145719 – 146159	146	ATG	hypothetical protein				
230	146447 – 146971	174	ATG	gp53: baseplate wedge subunit	baseplate wedge subunit	8.00E-35	Cr30	YP_009098986.1
231	146964 – 148244	426	ATG	hypothetical protein	hypothetical protein	4.00E-13	RSP15	BAU40030.1
232	148257 – 149174	305	ATG	gp5: baseplate hub and tail lysozyme	baseplate hub and tail lysozyme	5.00E-45	uvMED	BAR35966.1
233	149207 – 149608	133	ATG	gp25: baseplate wedge subunit	baseplate wedge subunit	5.00E-30	uvMED	BAR29104.1
234	149610 – 151400	596	ATG	gp6: baseplate wedge protein	baseplate wedge subunit	2.00E-122	Cr30	YP_009098969.1
235	151407 – 155273	1288	ATG	gp7: baseplate wedge subunit protein	baseplate wedge subunit	4.00E-62	Cr30	YP_009098968.1
236	155314 – 157041	575	ATG	gp8: baseplate wedge subunit protein	baseplate wedge subunit	6.00E-44	ΦM12	YP_009143024.1
237	157096 – 162570	1824	ATG	VlrC protein	hypothetical protein	1.00E-122	Cr30	YP_009098966.1
238	162570 – 164237	555	ATG	VlrC protein	VlrC protein	2.00E-34	uvMED	BAR36370.1
239	164242 – 164460	72	ATG	hypothetical protein		8.00E-36		
240	164457 – 165479	340	ATG	hypothetical protein	hypothetical protein	2.00E-14	uvMED	BAQ90538.1
241	165469 – 165816	115	GTG	hypothetical protein	hypothetical protein	1.00E-04	MED4-213	YP_007673777.1

[a] gene encoded on reverse strand. Abbreviations: *Moumouvirus goulette, Alistipes inops, Janthinobacterium lividum, Thioalkalimicrobium aerophilum, Pseudomonas stutzeri, Pseudomonas mendocina, Methylomonas denitrificans, Neptuniibacter caesariensis, Bradyrhizobium retamae, Flavobacterium filum, Hyphomonas adhaerens, Clostridium botulinum*.

Figure 6-3: Predicted promoter sequence in DLP6. Putative phage promoter consensus sequence was identified using PHIRE and plotted in WebLogo 3.

Rho-independent transcription termination sites were discovered using ARNold, which utilizes two complementary programs, Erpin[270] and RNAmotif [271]. Of the 54 potential termination sites identified, only 19 were retained as authentic because they had a ΔG of -10 kcal/mol or less (Figure 6-3 and Table 6-4). The 19 terminators were found downstream of AAY80_018 (precursor of major head subunit), AAY80_042 (triphosphate pyrophosphohydrolase), AAY80_057 (hypothetical protein), AAY80_090 (hypothetical protein), AAY80_091 (RuvC), AAY80_108 (hypothetical protein), AAY80_120 (hypothetical protein), AAY80_138 (hypothetical protein), AAY80_166 (hypothetical protein), AAY80_168 (hypothetical protein), AAY80_175 (endolysin), AAY80_180 (hypothetical protein), AAY80_199 and AAY80_200 (hypothetical proteins), AAY80_202 (ssDNA binding protein), AAY80_203 (lysozyme), AAY80_208 (endonuclease protein), AAY80_209 (hypothetical protein) and AAY80_212 (baseplate tail tube cap protein).

Determination of DLP6 phylogeny using the portal protein gp20

A BLASTn search using DLP6 genome as a query revealed the phages most similar to DLP6 are the *Sinorhizobium* phages ΦN3, ΦM19, ΦM7, and ΦM12. All four phages have coverage of only 4% with a 72% identity. Although initial BLASTp and BLASTn searches indicated DLP6 was more closely related to members of the T4-superfamily of phages, more comparisons were required to classify DLP6 as a T4-superfamily phage. For a commonly used phylogenetic comparison, the protein sequence of portal protein gp20 was used in a BLASTp search to identify 250 gp20 proteins[282-284]. The most significant results came from cyanophages grouped into the T4-superfamily. A MUSCLE alignment was completed using the 250 BLASTp results compared to the DLP6 gp20 protein (AAY80_011). This alignment was then used to generate an unrooted tree with the FastTree plugin for Geneious (Figure 6-4). The generated tree positions DLP6 in a clade with *Sinorhizobium* phages ΦM12, ΦN3, and *Caulobacter* phage Cr30 (Figure 6-4).

Table 6-4: Predicted Rho-independent terminators in DLP6 identified by the ARNold program. Putative terminators with a ΔG value of -10 kcal/mol or less were retained. DNA that is predicted to form the loop in the RNA is emboldened, whereas DNA that is predicted to encode an RNA stem is underlined.

Start	Program	Strand	Sequence	-ΔG
15290	both	+	CAATAAGAGAAGCCGCC**GCAA**GGCGGCTTTT	16.80
32949	both	+	AAGCCAGACAAGCCCCAGG**CTCCGC**CCTGGGGCTTTT	17.50
43682	both	+	TCAACGACTTAGCCCCAGA**CCCCGT**TCTGGGGCTTTT	16.10
72171	both	+	CTAATTGGAAAGCCGCCTCCGGGCGGCTTT	13.40
72816	Rnamotif	+	GACTTAGCGAAGCCCCC**GCCT**GGGGGCTTT	13.50
80627	both	+	TTGATGGAAAAGGCTCTCT**GCGGCTAACCGGCTC**AGAGAGCCTTT	13.46
85964	both	+	GACTTAGCGAAGCCCCGC**CCT**GCGGGGCTTT	14.80
93801	both	+	CGCCCCTGGAAGCCCGCTT**GGTCCGAGTGACT**AGGCGGGCTTTT	13.32
112696	both	+	GTAATCCAAAAGGGCTGGT**GTCCAAGAT**GCCAGCCCTTTT	15.70
113321	Rnamotif	+	CTATCTCGAAAGCCGCC**GCAA**GGCGGCTTT	16.80
117336	Rnamotif	-	ATCCATCTCGTCGAGAAGC**AGGTCTTCCT**GCTTCTCGTTTT	11.70
119975	both	+	TCAACGACTTAGCCCCTGA**GCCACC**TCAGGGGCTTTTCATTCCTG	17.10
128844	both	+	CTGTGAGAAAAGCCCTGCC**TTGATC**GGCGGGGCTTTTCCTTTGAT	17.60
129411	Erpin	+	CTTGACCGAAAAGCCCCGAAAGGGGCTTTTCTTTTGCCCA	14.90
131176	both	-	GCAAAGAAAAAGCCCAGGC**ATTGC**GCCTGGGCTTTTCAATTACAT	18.20
131178	both	+	GTAATTGAAAAGCCCAGGC**GCAAT**GCCTGGGCTTTTTCTTTGCG	17.80
136291	both	-	GATACAAGAAAGGCTCCC**TCT**CGGGAGCCTTTTGCTTTCACT	15.00
136293	both	+	TGAAAGCAAAAGGCTCCC**GAGA**GGGAGCCTTTCTTGTATCA	17.80
139414	Rnamotif	+	GACTTAGCGAAGCCCCC**GAAA**GGGGGCTTTACTTTTGGG	17.60

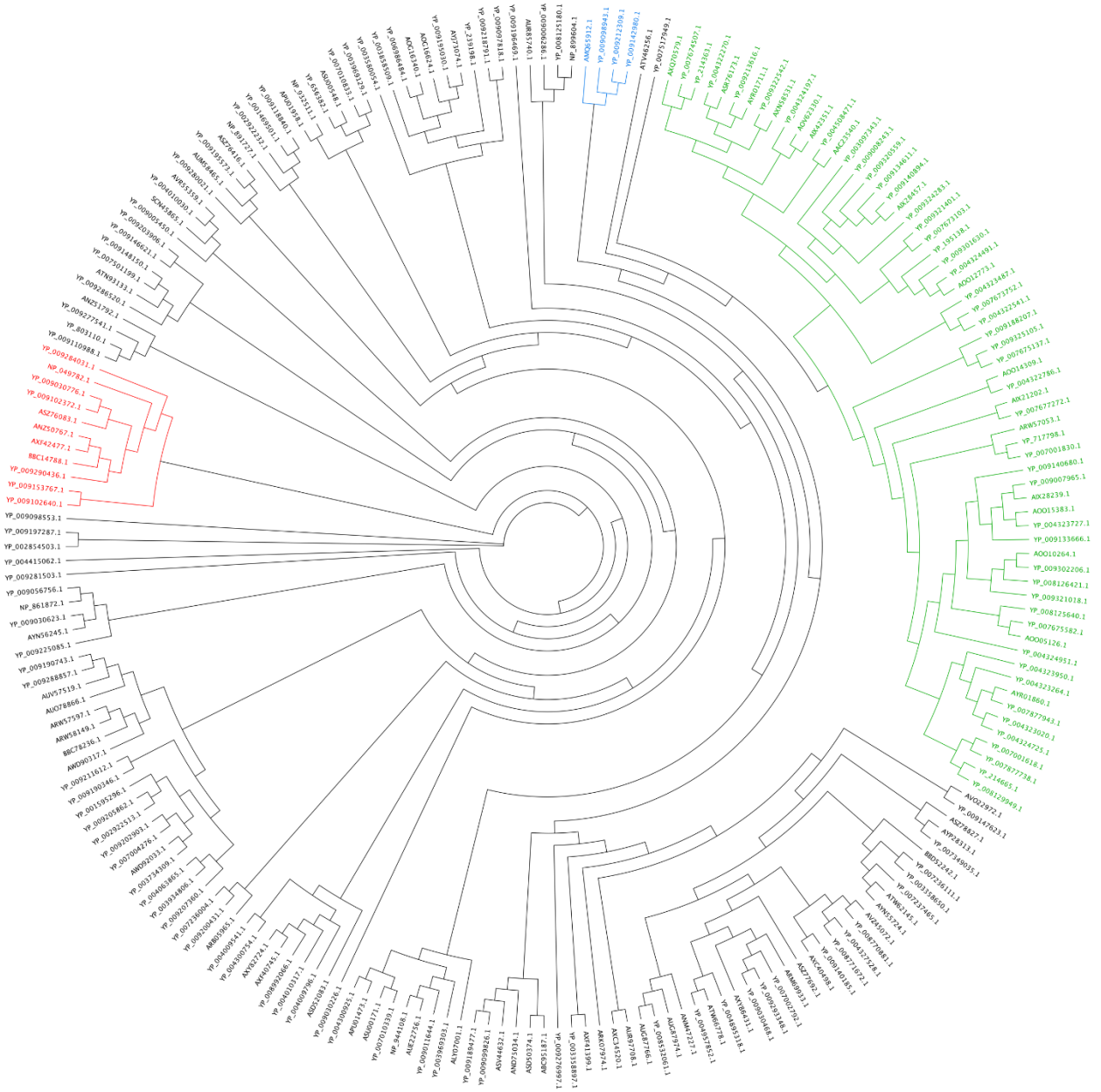

Figure 6-4: Unrooted gp20 (portal vertex protein: AAY80_011) tree. FastTree was used to generate the tree from a MUSCLE alignment between DLP6 gp20 and the top 216 BLASTp sequence results from the NCBI virus database. Clades of interest are marked as follows: DLP6; blue, T4; red, T4-like cyanophages; green.

Table 6-5: MUSCLE alignment percent identity score of full-length protein sequences of DLP6 against universal core and nearly universal core proteins of 18 T4-superfamily phages. Numbers indicate percent identity to the related DLP6 protein.

Gene product: function	Cyanophages												Enteric phages					
	ΦM 12	**S-SM2**	**P-SSM 4**	**P-Syn 1**	**HTVC008M**	**S-Syn 9**	**S-Syn 19**	**S-ShM 2**	**P-Syn 33**	**S-SM1**	**P-HM 1**	**S-CRM01**	**T4**	**Aeh**	**44RR**	**KVP 40**	**RB43**	**ΦW-14**
gp13: neck protein	53.2	35.5	34.2	34.3	39.1	36.1	35.7	33.5	35.7	37.2	36.1	22.9	21.7	19.4	20.2	23.1	25.2	22.6
gp14: neck protein	43.9	17	28	28.8	26.3	24.4	27.2	23.2	22.8	28.6	18.3	20.4	24.7	26.7	24.8	27.7	24.5	24.8
gp15: proximal tail sheath stabilization	40	31.9	27.9	32	31.5	33.2	27.3	34.8	34.1	33.3	27	30.9	24.3	26.8	23.5	16.6	23.6	20.8
gp16: small terminase	30.7	27.6	25.9	25.5	21.4	29.2	27.5	31.1	26.7	30.4	30.4	30.1	21.7	18.6	16.7	17.6	23.2	15.2
gp17: large terminase	42.5	47.2	48.1	47.2	46.9	48.5	48.7	48.9	49	48.8	46.1	42.8	27.8	28.5	27.7	29.8	29.8	32
gp18: tail sheath monomer	48.4	33.3	37.3	35	43.7	35.7	38.1	13.5	37.6	38.8	34.4	32.8	31.7	31.8	30.9	31.5	31.8	21.9
gp19: tail tube monomer	46.8	40.1	33.5	34.2	32.6	33.5	35.3	28.2	33.7	34	30.1	40.7	33.3	30.6	31.8	36.8	34.9	13.8
gp20: protal vertex protein	46.1	46.7	47.7	46.6	43.2	49.1	49.1	46.7	46.7	49.6	46.2	40.7	33.4	32	33.7	36	32.7	31.2
gp21: prohead core scaffold/protease	41.3	49.3	47.7	54.1	51.6	49.1	50	49.1	47.7	49.1	49.5	51.1	32.1	32.5	32.8	34.6	32.4	30
gp22: scaffold prohead core protein	22.3	29.1	26.5	29.9	30.1	26.5	26.5	26.6	27	25.9	27.8	25.9	19.3	20.2	22.1	17.8	20.6	11.6

gp23: precursor of major head subunit	65.4	59.6	55.7	52.5	60.5	57.1	58	56.3	58	55.8	55.9	53.8	34.4	33.3	34.8	35	34.3	44.8
DexA: exonuclease		32.5	32.1	35.7	30.5		33.3	35	34.2	32.9	29.3	22.2	12.2	13.7	11.2	8.2	14.6	11.4
gp3: head proximal tip of tail tube	27.3	31.5	29.1	31.4	18.5	29.3	27.6	28.3	29.2	30.9	29.4	27.2	20	18.5	17.8	17.5	20.3	14.8
UsvY: recombination, repair and ssDNA binding	39.6	27.1	30	22.4	29.2	27.9	27.2	31.9	26.6	27.7	28.5	27.8	20.8	18.1		20.8	16.7	19.7
UsvW: helicase	40	42.9	44.8	41.1	47.9	47.1	44.8	47.2	45.8	45.5	36.8	39.5	32.3	33.9	32.1	36	31.4	30.7
gp55: sigma factor, late transcription	36.6	37.3	43.5	42.5	41.6	40.7	40.4	44.1	38.4	42.3	42.5	37.8	23.9	24.8	22	25.4	22.8	24.7
gp47: recombination endonuclease subunit	42.7	33.9	36.7	34.3	39.4	39.2	40.3	36.2	35.5	39.8	37.7	38.5	27.5	23.1	24.9	27.1	28.2	22.4
gp46: recombination endonuclease subunit	45.7	42.9	43.4	43.2	44.3	44.3	45.9	45	44.9	45.2	44.2	37.8	31.2	25.5	29.2	28	31.3	20.5
gp45: sliding clamp accessory protein	34.2	43.8	40.1	38.6	42.2	33.5	40.8	35.9	38.8	39.6	40.6	39	28	27	26.9	36.9	26.1	25.8
gp44: clamp loader subunit	46.8	55.7	54.1	48.9	52.9	56.2	54.3	55.3	54.6	55.3	52.4	48.1	34.4	36.6	35.3	35.7	32.5	32.8

gp6: clamp loader subunit	37.1	31.8	29.6	32.1	35.6	33.1	31.6	42.5	33.8	31.6	37.2	31.8	21.1	18.1	19.5	27.3	21	25.5
RegA: early gene translational repressor	52.2	60	60	59.2	59.2	60	62.3	60.8	60	60.8	60	56.2	49.6	48.4	48.4	48.4	41.6	33.1
gp43: DNA polymerase	39.7	42.2	41.1	40	38	42.6	41.1	40.9	41	42.6	41.9	39.6	28.7	31.3	18.8	31.3	29.8	26.4
UsvX: recombination protein	47.9	23.9	25.3	26.4	25.1	25.3	25.3	24.5	24.7	26.1	25.8	26.2	49.9	52.7		47.8	46.1	35.1
gp41: DNA primase-helicase	50.5	57	58.8	53.7	54.6	53.5	53.7	56.2	53.5	54.3	53.8	51.7	37.4	40.9	39.7	38.6	38.7	31
Td: thymidylate synthase	11.4	12.5	11.3	12.5		10.9	14.4	12.3	11.6	11.9	11.4	10.6	11.4	9.3	10.8	7.4	12.1	8.8
gp61: DNA primase	44.6	34	35.8	37.8	37	39.8	36.2	37.5	39.2	40.8	39	37.5	31.1	32.4	33.4	31.6	30.9	30.6
NrdA: ribonucleotide reductase subunit A		46.8	46.3	47.4	45.5	47.7	47.3	46.5	48.1	47.9	47.9	48.1	43.6	29.9	17	44.3	43.8	13.1
NrdB: ribonucleotide reductase subunit B		44.2	45.9	42.6	44.6	45.1	44	44.5	41.2	44.5	44	41.2	37.5	39.3	18.3	41.7	37.7	19.4
gp33: late promoter transcription factor	18.7	23.3	22.7	27.4	21.3	23.8	24.4	22.6	22.4	26.2	29.9	26.7	22.7	19	24.1	20.7	14.8	30.2
NrdC: glutaredoxin	39	29.5	29.5	30.9	32.2	25.4	30.8	29.5	28.2	28.2	25.6	29.1	36.8	27.3		25.3	34.4	
gp59: DNA helicase loader	34.6	32.5	35.5		29.1	34.1	35.2	28.4	35.5	35	31.2	30.1	23.4	24.8	19.3	21.7	20.5	

gp32: ssDNA binding protein	48.8	48.3	46.7	44.5	44.7	46.5	47.1	48	49.5	48.5	47.3	46.7	34.2	30.2	30.6	33.1	25.3	
gp51: baseplate hub assembly catalyst	39.6	49.1	37.3	34.9	28.9	33.3	38.6	48.3	31.8	35.6	39.6	31	7.1		8.1	22.7	5.5	9.7
gp26: baseplate hub	36.3	31.1	34.3	32.1	36.1	35.4	36.3	34.4	35.6	38.9	31	27.2	19.1	23	15.8	24.6	14.8	24.3
gp4: head completion protein	50.3	46.3	45.3	35.1	38.3	47.3	46.6	44.6	45.9	45.3	46.6	30	38.2	38.2	47.4	39.2	37.9	36.9
gp48: baseplate tail tube cap	19	14.6	17	15.6	15.9	17.4	12.9	19.8	19.8	15.5	12.8	15.8	15.5	14.7	14.2	13.1	11.1	22.8
gp53: baseplate wedge	29.8	21	20.9	17	23.6	23.5	20.1	23.8	24.4	22.4	15.5	21.8	16.1	19.7	20.8	20.2	18	20.9
gp5: baseplate hub and tail lysozyme	11.2	9.4	9.6	5.8	22.2	8.7	8.7	9.2	9	8.6	7.2	8.7	16.7	15.9	15.7	22.3	13.8	10.3
gp25: baseplate wedge	44	27.8	35.1	33.8	36.6	30.8	29.1	28.4	29.9	26.1	27.1	30.1	25.8	27.9	25.6	27.3	28.8	25.2
gp6: baseplate wedge	33.1	26.7	24.3	30.5	31.6	24.7	24.7	23.1	25.1	23.5	29.6	27.3	21.5	21.6	23.6	21.8	22.7	22.6
gp7: baseplate wedge initiator	11.3	5.3	4	6.3	7.4	4.7	4.5	4.4	4.3	4.7	4.8	4.4	11.1	10.8	10.5	10.1	11.1	
gp8: baseplate wedge	20.1	16.4	19.6	14.4	18.3	19.3	19.7	19.7	19.1	20.3	17.4	18.4	10.5	12.1	11	10.9	11.5	

DLP6 contains features from T4-superfamily enteric and cyanobacteria phages

Genomic organization of DLP6 is similar to the *Sinorhizobium* transducing phage ΦM12, which has been classified into a new T4-superfamily fusing features of cyanophages and phages of enteric bacteria [275]. A set of "core universal" and "nearly universal" proteins has been determined for T4-superfamily phages [284,285]. DLP6 contains all of the core and nearly universal proteins common to all T4-superfamily phages (Table 6-5). A MUSCLE comparison was used to align the T4-superfamily phage core proteins against their respective orthologs from 18 T4-superfamily phages, and their percent identity to each ortholog was determined. The order of the gene products presented in Table 6-5 corresponds to the order they are found within the DLP6 genome. This order differs from the T4-superfamily phages, which start with gp41 (DNA primase-helicase). Results from the MUSCLE alignment reveal that DLP6 core proteins share the highest identity to their orthologs from cyanophages, except for the UvsX protein, which is most similar to T4-superfamily phages of enteric bacteria. The alignment also shows DLP6 has the highest percent identity to ΦM12 core proteins. The ortholog with the highest identity to a DLP6 protein was ΦM12 gp23, with 65% identity. Overall, the most highly conserved proteins between DLP6 and the T4-superfamily cyanophages are RegA (early transcriptional regulator) and gp23 (major capsid protein), with identity averaging 59 and 57% respectively. This suggests that DLP6 is divergent from the T4-superfamily phages. DLP6 does share additional proteins that are found within the T4-superfamily cyanophages or the T4-superfamily enteric phages.

All sequenced T4-superfamily cyanophages feature an accessory core set of 25 gene clusters (T4-GCs) that are not found within enteric bacteria T4-superfamily phages[284]. Of this accessory core, DLP6 encodes six of the core proteins: CobS (porphyrin biosynthetic protein), PhoH (P-starvation inducible protein), T4-CG 313 (hypothetical protein), T4-GC 321 (hypothetical protein), VlrC (predicted structural protein) and MazG (pyrophosphatase) (Table 6-6). Although DLP6 does contain these T4-superfamily cyanophage accessory core proteins, the DLP6 proteins are again divergent, with the maximum identity found with the PhoH (P-starvation inducible) protein at 45% to P-HM1 cyanophage.

Table 6-6: MUSCLE alignment percent identity score of DLP6 amino acid sequences against T4-superfamily cyanophages accessory core proteins. DLP6 contains six of the 25 T4-superfamily cyanophage core proteins. Numbers indicate percent identity to the related DLP6 protein.

Gene product: function	**Cyanophages**											
	ΦM 12	**S-SM2**	**P-SSM 4**	**Syn1**	**HTV C00 8M**	**Syn9**	**Syn1 9**	**S-ShM 2**	**Syn3 3**	**S-SM1**	**P-HM 1**	**S-CR M01**
CobS: porphyrin biosyntheic protein		41.6	41.9	43.6	40.3	42.2	41.2	41.9	41.5	41.1	42.7	40.7
PhoH: P-starvation inducible protein	33.0	42.5	44.4	44.7	42.8	44.0	43.8	42.2	43.0	42.6	44.8	39.2
T4 Gc 313: hypothetical protein	33.5	30.6	31.2	36.0	33.1	30.5	31.9	31.4	30.5	31.9	26.4	28.1
T4 Gc 321: hypothetical protein	39.2	25.3	17.7	20.7	30.4	24.1	25.3	25.3	17.7	24.1	21.0	14.3
VlrC: predicted structural protein	19.3	14.7	16.3	13.0	17.5	16.2	16.4	15.5	16.3	15.8	14.9	13.7
MazG: pyrophosphatase		16.9	16.4	20.6		16.3	19.4	16.9	16.3	18.1	16.9	16.3

DLP6 was found to contain nine out of the designated 30 non-cyanophage core T4-superfamily proteins. A MUSCLE alignment of these nine proteins indicates that although this phage does contain the proteins, they do not share high amino acid identities (Table 6-7). The protein sharing the highest identity was the Rnase H protein, sharing 32.6% identity with the RB43 phage protein. Again, these results demonstrate that DLP6, although classified within the T4-superfamily, it is more divergent than the other accepted members.

Differences between the T4-superfamily superfamily members discussed in this paper as compared to DLP6 are interesting, given that DLP6 contains the complete set of the T4-superfamily core group of proteins, six accessory core cyanophage proteins, and nine non-cyanophage accessory core proteins. Moreover, the ends of DLP6 genome feature 229 bp direct

terminal repeats, unlike the genomes of other members of the T4-superfamily of phages which are circularly permuted. The direct terminal repeats are located in a region of DNA devoid of ORFs and has no identity to known DNA sequences in the NCBI database. The average number of tRNAs encoded by the T4-superfamily phages is ten, except S-CRM01, a freshwater cyanophage that encodes 33 tRNAs[286]. The DLP6 genome contains 30 tRNAs, which is in the high range for the T4-superfamily. DLP6 also encodes a transposase (AAY80_133) (Figure 6-2 and Table 6-7).

Table 6-7: MUSCLE alignment percent identity score of DLP6 amino acid sequences against T4-superfaimly non-cyanophage core proteins. This table includes nine of the 32 non-cyanophage core proteins. Numbers indicate percent identity to the related DLP6 protein.

Gene product: function	Enteric phages					
	T4	**Aeh**	**44 RR**	**KVP 40**	**RB43**	**ΦW-14**
RnaseH	28.9	32.5	28.9	32.2	32.6	8.8
gp1: dNMP kinase	12.9	10.8	10	12.3	8.9	9.2
dCMP deaminase	24.5	24.6	22.8	26.6	18	16
DbsA: ssDNA binding protein	28.9	20.5	20.5	22.9	25.3	
gp34: long tail fiber	11.9	11.6	12.1	11.6	11.9	
gp30: dna ligase	24.7	23.5	21.9	29.7	24.5	32.3
Tk: thymidine kinase	13.5	13.5	12	11.4	13.5	8.4
gp31: head assembly cochaperone with GroEL	16.7	12.8	14	20	7.6	
gp59: loader of gp41 DNA helicase	23.4	24.8	19.3	21.7	20.5	

CONCLUSION

Stenotrophomonas maltophilia bacteriophage DLP6 was isolated from a soil sample using strain D1571 as the host bacterium. Phage DLP6 exhibits a moderate host range, infecting 13 out of 27 clinical *S. maltophilia* strains. A phylogenetic comparison of the gp20 portal protein against the top 250 BLASTp results places DLP6 in a clade with *Sinorhizobium* phages ΦM12, ΦN3, and *Caulobacter* phage Cr30. Although DLP6 does encode all of the T4-superfamily core and nearly universal core orthologs, the identity between these proteins and their nearest neighbors is typically less than 60%. The DLP6 genome also encodes six T4-superfamily cyanophage core proteins, but again, the nearest neighbor identity is below 40%. There are nine

T4-superfamily non-cyanophage core proteins found within the DLP6 genome, though the identity between the DLP6 proteins and the T4-superfamily enteric phage orthologs averaged less than 30%. Although DLP6 also encodes a transposon, experimental investigation suggests it does not form a stable prophage in *S. maltophilia* strain D1571. The results presented in this paper suggest DLP6 is a divergent T4-superfamily phage, exhibiting characteristics not yet identified in other T4-superfamily phages.

ACKNOWLEDGEMENTS

The authors would like to thank David Speert of the Canadian *Burkholderia cepacia* complex Research and Referral Repository (CBCCRRR) for the kind gift of numerous *S. maltophilia* strains. The authors would also like to thank the Provincial Laboratory for Public Health— North (Microbiology), Alberta Health Services, for access to *S. maltophilia* clinical isolates.

Chapter 7

Conclusions and future directions

Research goals

Members of the Smc are significant opportunistic pathogens against immunocompromised individuals and treatment of these infections is exceedingly difficult due to their high levels of innate antibiotic resistance and high mutation rate in established infections[64,70,119,128]. Transmission of this pathogen in hospital environments is often due to contaminated cleaning solutions and water sources[24]; though, research has also shown viable *S. maltophilia* can be recovered from cough-generated aerosols of cystic fibrosis patients[60,61]. The high levels of innate antibiotic resistance in both environmental and clinical isolates are due to multiple mechanisms such as chromosomally encoded beta-lactamases[101], aminoglycoside-modifying enzymes [104,130], a quinolone resistance gene[98], and at least ten efflux pumps[287]. The current drug of choice against *S. maltophilia* infections is TMP/SMX, though resistance to this drug is also on the rise globally[114]. The high levels of innate antibiotic resistance are problematic, as patients who receive the wrong antibiotic from the start of treatment are shown to have high rates of mortality[55].

With a slowdown in the production of novel antibiotics, alternative therapeutics must be investigated to combat *S. maltophilia* infections. A promising strategy to fight these extremely antibiotic resistant infections is the use of bacteriophages to selectively kill the pathogenic bacteria. A rise in antibiotic resistance in many bacterial species globally has caused a renewed interest in the clinical application of phages. A total of 12 clinical trials have been documented, with the majority of the completed trials showing positive results[140,141,143]. Additionally, the first phage therapy center was opened in North America called the Center for Innovative Phage Applications and Therapeutics (IPATH). The opening of this treatment center highlights the renewed phage therapy interest in the West.

Phage therapy treatment of *S. maltophilia* infections first requires the characterization of lytic phages. To this end, six phages were isolated from Edmonton, Canada soil samples using clinical *S. maltophilia* strains D1585, D1571, and D1614. Phages DLP1 and DLP2, detailed in Chapter 2, were isolated from two separate sampling sites against strain D1585[163]. DLP1 and DLP2 are closely related, sharing 96.7 % identity over 97.2 % of their genomes. Although so closely related genomically, the phages exhibit differing host ranges within *S. maltophilia* strains and they both have an unusually broad host range, capable of infecting *P. aeruginosa* isolates.

These phages are promising for phage therapy as *S. maltophilia* and *P. aeruginosa* co-infections have been shown to cause increased mortality in patients[24].

The T4-like phage DLP6, discussed in Chapter 6, was isolated using *S. maltophilia* strain D1571 and shows a moderate host range within the 27 clinical isolates by infecting 13 strains. DLP6 is the only *Myoviridae* phage our lab has isolated against *S. maltophilia* to date. Although DLP6 was found to encode a transposase, it is unable to lysogenize the host strain D1571[164]. This phage would be beneficial to include in a phage cocktail against *Stenotrophomonas* due to its moderate host range and lytic lifestyle.

Characterization of the three remaining phages detailed in Chapters 3-5 revealed DLP3-DLP5 are temperate phages capable of lysogenic conversion of their hosts. Chapter 4 details the characterization of phage DLP4 which was found to be capable of stable lysogeny in spite of the lack of an identifiable repressor, integrase, or transposase. DLP4 was found to cause the lysogenic conversion of its host strain D1585 though the expression of *folA.* The DLP4 *folA* results in increased TP resistance of the lysogen compared to wild-type D1585 at a statistically significant level. DLP4 also encodes the virulence factor *ybiA* which is responsible for increasing swarming motility in *Escherichia coli*. This gene is expressed during the lysogenic cycle, though no phenotypic change were observed in the lysogen due to this gene.

In Chapter 5, the establishment of DLP5 as the type species for the new genus *Delepquintavirus* is detailed. This phage is capable of lysogenizing its host strain D1614 as a phagemid and is found to encode cell wall modifying enzymes. DLP5 also causes an increase in growth rate during the lag and early exponential phase of lysogen growth. Additionally, chapter 3 details phage DLP3 which is identified as the second member of the *Delepquintavirus* genus. DLP3 is also capable of stably lysogenizing its host. Investigation into a hypothetical protein encoded by DLP3 led to the discovery of an erythromycin resistance protein which causes a statistically significant increase in erythromycin resistance in the host during lysogeny.

Aside from the assessment of possible phage therapy candidates, the data presented here provides vital insights into *S. maltophilia* phages. Only 20 phages have been isolated and characterized against *S. maltophilia* to date, including the six phages discussed in this thesis. Thus, each novel *Stenotrophomonas* phage documented significantly increases our understanding of these phages in their diversity, relatedness, and distribution. For example, the isolation of such closely related phages as DLP1 and DLP2 from separate sites in the Edmonton area was quite

surprising. With further analysis of these phages, we discovered they are also closely related to *P. aeruginosa* phages PK25, PA73, and AB26. This discovery led to the extended host range screen with *P. aeruginosa* isolates which uncovered the unique broad host ranges of DLP1 and DLP2. The results challenge previous ideas held about phage biology and what broad and narrow tropism actually means with respect to phages.

Additionally, while characterizing the DLP5 phage, it was evident this phage was divergent from any phage isolated to date due to the limited identity found in the NCBI database. This led to the establishment of the new genus *Delepquintavirus,* which DLP3 was later included in due to its genetic identity to DLP5. Both phages were found to encode two ParB proteins typically used in plasmid partitioning, but the ParA protein has not yet been identified. Characterization of these two phages led to the identification of a new erythromycin resistance protein, and potentially, a new type of partitioning system which has not yet been characterized. Like DLP5, the limited identity of DLP6 to any T4-like phages in the NCBI database is also interesting as the T4-like *Stenotrophomonas* phage Smp14[155] does not share any nucleotide identity to DLP6. This highlights the extreme diversity which can be found within the T4-like group of phages. These examples highlight the need for the isolation and characterization of additional *Stenotrophomonas* phages to enable comparative genomic analyses and the identification of phages for use in therapy against this pathogen.

Future directions

Lytic phages

Infections caused by *S. maltophilia* are great candidates for phage therapy due to the treatment difficulty experienced as a result of high levels of antibiotic resistance. In order to generate optimal phage cocktails against this pathogen, additional lytic phages must be isolated and characterized to expand our phage library. The easiest method to isolate phages against *S. maltophilia* continues to be through the use of soil samples. Hospital sources could also be used, but this would come with greater safety risks compared to soil samples due to the possibilities of other infectious agents contaminating the samples. For a more targeted approach to phage hunting, mutants resistant to the current lab phages can be isolated and used. This approach would help isolate phages which ‘guard’ against cocktail resistance due to receptor mutations of

the target bacteria. For example, the type IV pilus (T4P) has been identified as the potential receptor for many of the DLP phages characterized to date. T4P mutants of the lab *S. maltophilia* strains could be used in phage hunting experiments to enrich for phages which use an alternative receptor. This is important as phage cocktails should ideally have a mixture of phages which use different receptors; thus, identifying new phages which do not use the type IV pilus is crucial.

Further optimization of DLP1, DLP2, and DLP6 should be completed before they are included in therapy. The phages can be repeatedly passaged on the strains of *P. aeruginosa* (for DLP1 and DLP2) and *S. maltophilia* which they are capable of successfully infecting to increase their efficacy within their current host range. This directed evolution approach could also be applied to try and broaden the host ranges into strains that do not appear to be susceptible to phage infection. Successful infection of a bacterium is accomplished by injection of the phage genome into the host bacterium, propagation of the genome and expression of phage proteins which ultimately concludes with lysis of the host cell to release the phage progeny. If the phage is naturally virulent against the bacterium, this cycle would continue and could be observed as the appearance of plaques on a bacterial lawn. If a phage is capable of injecting its genome into the host cell, though not capable of successfully propagating in that host cell for a multitude of reasons, it would appear as though that bacterium is resistant to the phage. Bacterial lawns infected by that phage would not show visible plaques, though some escapee phage mutants may have been able to undergo successful infection resulting in lysis of bacteria at very low levels. Repeated passage of the phage on what appear to be non-susceptible hosts may lead to the isolation and propagation of escapee mutant variants capable of completing the lytic cycle within that host. These mutants could then be incorporated into the phage cocktail to expand its efficacy.

Pooling the trained phage stocks discussed above with the escapee mutant stocks would enable a broader host range within both *S. maltophilia* and *P. aeruginosa* for the phages without the need of molecular techniques. In combination with the directed evolution experiments, the receptor binding proteins of the phage could also be enhanced. This is particularly applicable for the phages DLP1 and DLP2 due to their high genomic identity but distinct host ranges. Identifying and deleting a non-essential gene in DLP1 and DLP2 would be beneficial to provide space on the phage genome to insert the receptor binding protein (RBP) of the other phage beside the original RBP to ensure properly timed expression. This would create a single genetically

modified hybrid phage with the host range of both DLP1 and DLP2 combined. By deleting a non-essential gene, the genome size would be similar to that of the wild type which is important for proper packaging of DNA into the phage capsid.

This process could be repeated using the receptor binding proteins of the closely related *P. aeruginosa* phages PA25, PA73, and Ab26 to expand the range of the hybrid phage into more *P. aeruginosa* strains. Although the phage stocks of those closely related *P. aeruginosa* phages have been lost, advancements in *de novo* DNA synthesis removes the need for those stocks as their genomic sequences are in the NCBI database. Identifying and deleting more non-essential genes within the hybrid genome would free up additional space to insert the genes encoding the receptor binding proteins of the PA25, PA73, and Ab26 *P. aeruginosa* phages. Of course, this process could be repeated with any other closely related phage identified in the NCBI database until there is no more space on the phage genome to accommodate the additional RBPs.

Genome engineering of the lytic phages could be accomplished using the well-known *Streptococcus pyogenes* CRISPR-Cas9 system as it has been shown to be an effective method for the genetic manipulation of lytic bacteriophages[289]. Lemay et al. (2017) described the efficient genetic manipulation of the Lactococcal phage p2 through deletions, insertions, and point mutations. The group demonstrated the ability to delete four genes with no deleterious effects to the lytic phage, suggesting those four hypothetical genes were not essential under the lab conditions studied. This system has not yet been optimized for use in *S. maltophilia;* therefore, this should be a top priority in the lab as it would be extremely useful for genetic manipulations of the *S. maltophilia* lab strains and the phages that infect them. After each deletion, the phage efficacy should be confirmed in the *Galleria mellonella* model to ensure the gene is not required during more physiologically relevant infection conditions than simply top agar overlays.

Optimization of the phage DLP6 could also be completed using the CRISPR-Cas9 system also. The use of CRISPR-Cas9 to edit the T4 genome has been explored and showed moderate success[290]. The main difficulty researchers faced in modifying the T4 genome was due to the DNA modifications of T4. Researchers found the restriction efficiency varied greatly depending upon the spacers chosen due to the modifications and so they were able to optimize the procedure by excluding sites with the adenine methylation sequence 5'-GATC-3'. DLP6 also has modified DNA; though, genetic manipulation of DLP6 could still be possible if the methylation sequences were also excluded from the spacers.

Further expanding on the DLP6 genome editing, deletion of the gene encoding the putative transposase would be a great target to ensure there is no chance of transposition within the bacterial hosts. DLP6 could also be modified to increase its lysis efficiency. Lysis inhibition is common to the T4-like phages and occurs when there is a higher concentration of phages compared to their bacterial host in a local environment. Lysis inhibition has been studied in T4 and is a complex phenotype that involves at least six T4 proteins[291]. This phenotype is not desirable in phages used for treatment as the main goal is to reduce the bacterial load in the patient as soon as possible. The genes involved in lysis inhibition in T4 have been identified; therefore, these genes should be easy to identify and delete from DLP6 genome.

Further genetic manipulations to delete the non-essential genes of DLP6 can be completed to enable 'space' on the DLP6 genome to insert additional RBPs from the other T4-like *S. maltophilia* phage Smp14 to expand the DLP6 host range. Luckily, the 16 kbp sequence submitted to NCBI encodes some of the structural proteins of Smp14; therefore, the relevant gene sequences could be submitted for *de novo* synthesis and ultimately used in the genetic manipulation of DLP6. This process could be repeated for any new T4-like phages identified against *S. maltophilia* if their sequence is submitted to NCBI. The genetic manipulations of DLP6 discussed above would enable the creation of a safer and more efficient phage for incorporation into the phage cocktail.

Once a robust *Stenotrophomonas* phage cocktail has been established with the optimized DLP1, DLP2, DLP6, and newly isolated 'guard' phages, *in vivo* studies should be conducted in *Galleria mellonella* to determine initial efficacy. These phages should all be tested alone and in various combinations against *S. maltophilia* strains. It should be mentioned that repeated attempts to use the *Galleria* model in the lab have proven to be exceedingly difficult due to the poor quality of wax worms we receive from the distributor as of late. Before the commencement of these experiments, establishing a better source for the wax worms is imperative so the studies can be conducted in a timely fashion. If the cocktails prove effective in the *G. mellonella* model, additional *in vivo* studies can then be conducted using murine models. The cocktail efficacy should be studied for the treatment of lung, skin, and bloodstream infections. This could be accomplished through aerosolization experiments for lung infections, topical applications for wound infections, and intravenous tail injections for bloodstream infections. Testing the optimized DLP1 and DLP2 phages against *P. aeruginosa,* and co-infections of *P. aeruginosa* and

S. maltophilia together in *G. mellonella* and murine animal models should be completed to see if they are efficacious. Co-infection rescue experiments are important as these bacteria are often isolated together from the lungs of CF patients.

The immunogenicity of each phage should be studied to determine what effect their administration may have on a patient. Basic studies could be completed using a mouse model to determine the immunogenic effect of the purified phage on mice. This could be accomplished with RNA transcriptome analysis to determine what changes occur in the expression of immunity-related genes involved in the pro- and anti-inflammatory response when the mice are exposed to the phages. Additionally, though a much more expensive option, could be the use of human peripheral blood monocytes (HPBM) to study the immunogenicity of each phage. The HPBM cells are precursors to tissue macrophages and dendritic cells and are central to immunological responses; therefore, using these cells would be ideal to determine the immune response to each phage within humans prior to clinical trials. Gene expression changes can be studied from the RNA transcriptome of each sample followed by a proteomic analysis see if there is a difference between the transcription and translation of the genes.

Further research into the receptor of DLP1 and DLP2, identified as the type IV pilus, can be conducted to gain insight into the host range of these phages. Glycosylation of the T4P was recently identified as an anti-phage strategy to prevent phage from binding and injecting their DNA into the target cell[292]. Harvey et al. (2018) discovered *P. aeruginosa* express one of five different T4Ps, two of which are glycosylated with polymers of *D*-arabinofuranose or O-antigen units. Both modifications were found to protect the *P. aeruginosa* cells from certain pilus-specific phages, as did alterations in the pilin sequences. Very little research has gone into the T4P of *S. maltophilia*; thus, this area of research has great potential to help elucidate what accounts for the host range variation observed between DLP1 and DLP2. The primary host of DLP1 and DLP2, *S. maltophilia* strain D1585, is also highly susceptible to many of the *S. maltophilia* phages isolated to date; thus, perhaps the glycosylation state of the D1585 T4P has some influence into its high susceptibility to many phages.

Temperate phages

Temperate phages play a major role in the bacterial acquisition of genes encoding virulence factors and antimicrobial resistance proteins. Due to this, further investigation into the

temperate phages our lab has isolated could provide greater insight into their role in virulence and antimicrobial resistance within clinical *S. maltophilia* isolates. Many annotated phage genomes still contain a multitude of genes encoding hypothetical proteins. Investigating the function of these proteins is important as it can help with the characterization of novel phages that are being investigated for use in therapy. For instance, due to the lack of identifiable proteins involved in phage lysogeny encoded on the DLP4 genome, one could assume DLP4 is a virulent phage. This is clearly not the case; therefore, it would be extremely beneficial to identify the proteins responsible for maintaining the lysogenic cycle of DLP4 within the host bacterium. The molecular methods used by DLP4 to maintain stable lysogeny would be interesting to elucidate as other phages could use a similar system to establish and maintain lysogeny in their bacterial host. This could help phage researchers quickly identify if a novel phage is temperate. Additionally, identifying the novel proteins involved in repressing the lytic cycle would greatly contribute to our general understanding of phage biology regarding the regulation of the lytic/lysogenic cycle in newly identified temperate phages.

To determine which proteins are responsible for the establishment of DLP4 lysogeny in the D1585 host, genes encoding hypothetical proteins could be sequentially deleted using the CRISPR-Cas9 approach discussed above. Deletion of the genes encoding hypothetical proteins with conserved domains of unknown function (DUF) should be targeted first using the CRISPR-Cas9 system. These genes are commonly found in many characterized phages, yet nothing is known about their function; therefore, deleting the genes could provide some insight into the importance of these genes to the various phages which encode them. Following that, sequential deletion of DLP4 genes encoding other hypothetical proteins should be completed in the effort to identify the DLP4 lytic cycle repressor as well as any essential genes. In addition to identifying and deleting the gene encoding the lytic cycle repressor, knocking out the *ybiA* and *folA* genes of DLP4 would create a genetically modified phage suitable for use in therapy as it would be obligately lytic and would no longer encode known virulence and antibiotic resistance genes. These experiments could also help advance phage biology through the identification of essential genes in closely related phages, while also possibly identifying the role of the ubiquitous DUF phage proteins.

Further research into any phenotypic changes of the lysogens should be studied to determine what effect these phages have on their host strains. Studying the lysogens and looking

for changes in virulence, stress tolerance, and antimicrobial susceptibility could provide some insight into phenotypic changes occurring due to phage lysogeny. All of the temperate phages isolated to date encode moron genes which are putative virulence factors as described in Chapters 3- 5. For example, DLP4 encodes a protein with a NeuB_NnaB family domain found in N-acetylneuraminate synthase proteins. The N-acetylneuraminate synthase protein produces N-acetylneuraminic acid (NANA), a sialic acid that is used by bacterial pathogens as food or to hide from their mammalian hosts[254]. When using NANA to hide from the human immune cells, the bacteria can either produce a sialic acid-containing capsule or sialylate their LPS. To determine if DLP4 is sialylating the D1585 LPS, the LPS could be isolated, and SDS-PAGE with silver staining could be conducted to see if there are migration differences between the lysogen and wild type LPS samples. Electrospray ionization-mass spectrometry could also be used for structural characterization the LPS samples.

The DLP3 and DLP5 phages encode putative cell wall modifying enzymes which could be expressed in the lysogen. To determine if the cell wall of the lysogen is modified by each phage, whole bacterial cell walls can be isolated, and their peptidoglycan digested. A composition analysis can then be conducted using Ultra Performance Liquid Chromatography (UPLC). This would provide detailed information such as the identity of the muropeptide components and their concentrations, the average length of the glycan strands, and the material involved in the crosslinks between the strands. If there are differences in the cell wall components between the lysogens and their wild type counterparts, this analysis will provide valuable information as to what types of modifications have occurred. Additionally, one could knock out the gene encoding the putative cell wall modifying enzyme in each phage and repeat the experiment to determine if that phage-encoded enzyme is responsible for the observed differences, or if the differences are due to altered gene expression in the host attributed to the presence of the prophage.

An investigation into potential virulence differences of the D1571::DLP3, D1585::DLP4, and D1614::DLP5 lysogens compared to their wild type strains can be conducted. The common duckweed (*Lemna minor*) model[293] can be employed to observe virulence differences between the strains quickly. This model can be used in low oxygen incubation chambers which mimic the conditions found within the mucus layer of the human lungs making the results physiologically relevant. The RNA transcriptome of the lysogen and wild type cells should be analyzed if

differences are observed with low oxygen conditions compared to ambient air. Further investigation into any changes in gene expression between the two strains could lead to interesting data on the prophage involvement in virulence at more physiologically relevant conditions. If differences are observed between the wild type strains and the lysogens, the study should be moved into cystic fibrosis lung cell lines at low oxygen conditions and the RNA transcriptome reanalyzed as the presence of human cells could greatly alter the results.

Significance

The isolation and characterization of bacteriophages against the extremely antibiotic resistant bacterium *Stenotrophomonas maltophilia* is vital due to an increase in infections caused by this organism and the high mortality rates observed in immunocompromised individuals. We have shown the ease with which *Stenotrophomonas* phages can be isolated from soil samples; thus, the limiting factor for building effective phage cocktails against this pathogen is the isolation of naturally lytic phages with broad host ranges. Half of the phages we have isolated to date are lytic with moderate host ranges, making them great candidates for inclusion in a phage cocktail. It is important to think about the temperate phages of *S. maltophilia* when isolating phage for therapy, as all of the phages characterized to date encode either antibiotic resistance genes, virulence factors, or both. The phage DLP4 is a perfect example of a temperate phage which may appear lytic due to the apparent absence of identifiable temperate genes; therefore, extensive characterization of each phage must be conducted to ensure the safety and efficacy of the therapeutic phages.

BIBLIOGRAPHY

1 Hugh, R. & Ryschenkow, E. *Pseudomonas maltophilia*, an *Alcaligenes*-like species. *J Gen Microbiol* **26**, 123-132, doi:10.1099/00221287-26-1-123 (1961).

2 Komagata, K., Yabuuchi, Y., Tamagawa, V. & Ohyama, A. *Pseudomonas melanogena* Iizuka and Komagata 1963, a later subjective synonym of *Pseudomonas maltophilia* Hugh and Ryschenkow 1960. *Int J Syst Bacteriol* **24**, 242-247 (1974).

3 Komagata, K. Differentiation of genus *Pseudomonas* and related aerobic bacteria. *Journal of General Applied Microbiology* **7** (1961).

4 Ulrich, J. A. & Needham, G. M. Differentiation of *Alcaligenes faecalis* from *Brucella* bronchisepticus by biochemical and nutritional methods. *J Bacteriol* **65**, 210-215 (1953).

5 Palleroni, N. J., Kunisawa, R., Contopoulou, R. & Doudoroff, M. Nucleic acid homologies in the genus *Pseudomonas*. *Int J Syst Bacteriol* **23**, 333-339 (1973).

6 Palleroni, N. J. & Bradbury, J. F. *Stenotrophomonas*, a new bacterial genus for *Xanthomonas maltophilia* (Hugh 1980) Swings et al. 1983. *Int J Syst Bacteriol* **43**, 606-609, doi:10.1099/00207713-43-3-606 (1993).

7 Wolf, A., Fritze, A., Hagemann, M. & Berg, G. *Stenotrophomonas rhizophila* sp. nov., a novel plant-associated bacterium with antifungal properties. *International Journal of Systematic and Evolutionary Microbiology* **52**, 1937-1944, doi:10.1099/ijs.0.02135-0 (2002).

8 Finkmann, W., Altendorf, K., Stackebrandt, E. & Lipski, A. Characterization of N2O-producing *Xanthomonas*-like isolates from biofilters as *Stenotrophomonas nitritireducens* sp. nov., *Luteimonas mephitis* gen. nov., sp. nov. and *Pseudoxanthomonas broegbernensis* gen. nov., sp. nov. . *Int. J. Syst. Evol. Microbiol.* **50**, 273–282 (2000).

9 Assih, E. A. *et al. Stenotrophomonas acidaminiphila* sp. nov., a strictly aerobic bacterium isolated from an upflow anaerobic sludge blanket (UASB) reactor. *International Journal*

of Systematic and Evolutionary Microbiology **52**, 559-568, doi:10.1099/00207713-52-2-559 (2002).

10 Heylen, K., Vanparys, B., Peirsegaele, F., Lebbe, L. & De Vos, P. *Stenotrophomonas terrae* sp. nov. and *Stenotrophomonas humi* sp. nov., two nitrate-reducing bacteria isolated from soil. *International Journal of Systematic and Evolutionary Microbiology* **57**, 2056-2061, doi:10.1099/ijs.0.65044-0 (2007).

11 Yang, H. C., Im, W. T., Kang, M. S., Shin, D. Y. & Lee, S. T. *Stenotrophomonas koreensis* sp. nov., isolated from compost in South Korea. *International Journal of Systematic and Evolutionary Microbiology* **56**, 81-84, doi:10.1099/ijs.0.63826-0 (2006).

12 Kaparullina, E., Doronina, N., Chistyakova, T. & Trotsenko, Y. *Stenotrophomonas chelatiphaga* sp. nov., a new aerobic EDTA-degrading bacterium. *Systematic and Applied Microbiology* **32**, 157-162, doi:10.1016/j.syapm.2008.12.003 (2009).

13 Sanchez-Castro, I. *et al. Stenotrophomonas bentonitica* sp. nov., isolated from bentonite formations. *Int J Syst Evol Microbiol* **67**, 2779-2786, doi:10.1099/ijsem.0.002016 (2017).

14 Lee, M. *et al. Stenotrophomonas daejeonensis* sp. nov., isolated from sewage. *Int J Syst Evol Microbiol* **61**, 598-604, doi:10.1099/ijs.0.017780-0 (2011).

15 Kim, H. B. *et al. Stenotrophomonas ginsengisoli* sp. nov., isolated from a ginseng field. *International Journal of Systematic and Evolutionary Microbiology* **60**, 1522-1526, doi:10.1099/ijs.0.014662-0 (2010).

16 Weber, M., Schünemann, W., Fuß, J., Kämpfer, P. & Lipski, A. *Stenotrophomonas lactitubi* sp. nov. and *Stenotrophomonas indicatrix* sp. nov., isolated from surfaces with food contact. *International Journal of Systematic and Evolutionary Microbiology* **68**, 1830-1838, doi:10.1099/ijsem.0.002732 (2018).

17 Yi, H., Srinivasan, S. & Kim, M. K. *Stenotrophomonas panacihumi* sp. nov., isolated from soil of a ginseng field. *Journal of Microbiology* **48**, 30-35, doi:10.1007/s12275-010-0006-0 (2010).

18 Ramos, P. L. *et al.* Screening for endophytic nitrogen-fixing bacteria in Brazilian sugar cane varieties used in organic farming and description of *Stenotrophomonas pavanii* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **61**, 926-931, doi:10.1099/ijs.0.019372-0 (2011).

19 Ouattara, A. S., Le Mer, J., Joseph, M. & Macarie, H. Transfer of *Pseudomonas pictorum* gray and thornton 1928 to genus *Stenotrophomonas* as *Stenotrophomonas pictorum* comb. Nov., and emended description of the genus *Stenotrophomonas*. *International Journal of Systematic and Evolutionary Microbiology* **67**, 1894-1900, doi:10.1099/ijsem.0.001880 (2017).

20 Handa, Y. *et al. Stenotrophomonas tumulicola* sp. nov., a major contaminant of the stone chamber interior in the Takamatsuzuka Tumulus. *International Journal of Systematic and Evolutionary Microbiology* **66**, 1119-1124, doi:10.1099/ijsem.0.000843 (2016).

21 Svensson-Stadler, L. A., Mihaylova, S. A. & Moore, E. R. *Stenotrophomonas* interspecies differentiation and identification by *gyrB* sequence analysis. *FEMS Microbiol Lett* **327**, 15-24, doi:10.1111/j.1574-6968.2011.02452.x (2012).

22 Rhee, J. Y. *et al.* Distinct groups and antimicrobial resistance of clinical *Stenotrophomonas maltophilia* complex isolates from Korea. *J Med Microbiol* **62**, 748-753, doi:10.1099/jmm.0.053355-0 (2013).

23 Ochoa-Sanchez, L. E. & Vinuesa, P. Evolutionary genetic analysis uncovers multiple species with distinct habitat preferences and antibiotic resistance phenotypes in the *Stenotrophomonas maltophilia* complex. *Front Microbiol* **8**, 1548, doi:10.3389/fmicb.2017.01548 (2017).

24 Brooke, J. S. *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clinical Microbiology Reviews* **25**, 2-41, doi:10.1128/cmr.00019-11 (2012).

25 Hugh, R. & Gilardi, G. L. in *Manual of clinical microbiology* (eds Lennette E H, Balows A, Hausler W Jr, & Truant J P) 288–317 (1980).

26 Mahdi, O., Eklund, B. & Fisher, N. in *Current Protocols in Microbiology* 6F.1.1-6F.1.6 (2014).

27 Martinez, P. *et al. Stenotrophomonas maltophilia* responds to exogenous AHL signals though the LuxR solo SmoR (Smlt1839). *Frontiers in Cellular and Infection Microbiology* **5** (2015).

28 Huang, T. P. & Lee Wong, A. C. Extracellular fatty acids facilitate flagella-independent translocation by *Stenotrophomonas maltophilia*. *Res Microbiol* **158**, 702-711, doi:10.1016/j.resmic.2007.09.002 (2007).

29 Nicoletti, M. *et al. Stenotrophomonas maltophilia* strains from cystic fibrosis patients: genomic variability and molecular characterization of some virulence determinants. *Int J Med Microbiol* **301**, 34-43, doi:10.1016/j.ijmm.2010.07.003 (2011).

30 Berg, G., Roskot, N. & Smalla, K. Genotypic and phenotypic relationships between clinical and environmental isolates of *Stenotrophomonas maltophilia*. *Journal of Clinical Microbiology* **37**, 3594-3600 (1999).

31 Chatelut, M., Dournes, J. L., Chabanon, G. & Marty, N. Epidemiological typing of *Stenotrophomonas* (*Xanthomonas*) *maltophilia* by PCR. *Journal of Clinical Microbiology* **33**, 912-914 (1995).

32 Hauben, L., Vauterin, L., Moore, E. R. B., Hoste, B. & Swings, J. Genomic diversity of the genus *Stenotrophomonas*. *International Journal of Systematic Bacteriology* **49**, 1749-1760, doi:10.1099/00207713-49-4-1749 (1999).

33 Minkwitz, A. & Berg, G. Comparison of antifungal activities and 16S ribosomal DNA sequences of clinical and environmental isolates of *Stenotrophomonas maltophilia*. *Journal of Clinical Microbiology* **39**, 139-145, doi:10.1128/JCM.39.1.139-145.2001 (2001).

34 Nesme, X., Vaneechoutte, M., Orso, S., Hoste, B. & Swings, J. Diversity and genetic relatedness within genera *Xanthomonas* and *Stenotrophomonas* using restriction

endonuclease site differences of PCR-amplified 16S rRNA gene. *Systematic and Applied Microbiology* **18**, 127-135, doi:10.1016/S0723-2020(11)80460-1 (1995).

35 Salah-Tazdaït, R. *et al.* Isolation and characterization of bacterial strains capable of growing on malathion and fenitrothion and the use of date syrup as an additional substrate. *International Journal of Environmental Studies* **75**, 466-483, doi:10.1080/00207233.2017.1380981 (2017).

36 Larik, I. A. *et al. Stenotrophomonas maltophilia* strain 5DMD: an efficient biosurfactant-producing bacterium for biodegradation of diesel oil and used engine oil. *International Journal of Environmental Science and Technology*, 1-10, doi:10.1007/s13762-018-1666-2 (2018).

37 Baldiris, R., Acosta-Tapia, N., Montes, A., Hernández, J. & Vivas-Reyes, R. Reduction of hexavalent chromium and detection of chromate reductase (ChR) in *Stenotrophomonas maltophilia*. *Molecules* **23**, doi:10.3390/molecules23020406 (2018).

38 Raman, N. M., Asokan, S., Shobana Sundari, N. & Ramasamy, S. Bioremediation of chomium(VI) by *Stenotrophomonas maltophilia* isolated from tannery effluent. *International Journal of Environmental Science and Technology* **15**, 207-216, doi:10.1007/s13762-017-1378-z (2018).

39 Rathi, M. & Nandabalan, Y. K. Copper-tolerant rhizosphere bacteria—characterization and assessment of plant growth promoting factors. *Environmental Science and Pollution Research* **24**, 9723-9733, doi:10.1007/s11356-017-8624-2 (2017).

40 Zhang, Y. F. *et al.* Characterization of ACC deaminase-producing endophytic bacteria isolated from copper-tolerant plants and their potential in promoting the growth and copper accumulation of *Brassica napus*. *Chemosphere* **83**, 57–62 (2011).

41 Banerjee, M. & Yesmin, L. Sulfur-oxidizing plant growth promoting rhizobacteria for enhanced canola performance. *Sulfur-oxidizing plant growth promoting Rhizobacteria for enhanced canola performance* (2002).

42 Ikemoto, S., Suzuki, K., Kaneko, T. & Komagata, K. Characterization of strains of *Pseudomonas maltophilia* which do not require methionine. *International Journal of Systematic Bacteriology* **30**, 437-447, doi:10.1099/00207713-30-2-437 (1980).

43 Park, M. *et al.* Isolation and characterization of diazotrophic growth promoting bacteria from rhizosphere of agricultural crops of Korea. *Microbiological Research* **160**, 127-133, doi:10.1016/j.micres.2004.10.003 (2005).

44 Ryan, R. P. *et al.* The versatility and adaptation of bacteria from the genus *Stenotrophomonas*. *Nat Rev Microbiol* **7**, 514-525, doi:10.1038/nrmicro2163 (2009).

45 Berg, G. & Martinez, J. L. Friends or foes: can we make a distinction between beneficial and harmful strains of the *Stenotrophomonas maltophilia* complex? *Front Microbiol* **6**, 241, doi:10.3389/fmicb.2015.00241 (2015).

46 Simon, S. & Petrášek, J. Why plants need more than one type of auxin. *Plant Science* **180**, 454-460, doi:**https://doi.org/10.1016/j.plantsci.2010.12.007** (2011).

47 Parapouli, M. *et al.* Molecular, biochemical and kinetic analysis of a novel, thermostable lipase (LipSm) from *Stenotrophomonas maltophilia* Psi-1, the first member of a new bacterial lipase family (XVIII). *Journal of Biological Research (Greece)* **25**, doi:10.1186/s40709-018-0074-6 (2018).

48 Verger, R. 'Interfacial activation' of lipases: facts and artifacts. *Trends in Biotechnology* **15**, 32-38, doi:**https://doi.org/10.1016/S0167-7799(96)10064-0** (1997).

49 Al-Anazi, K. A. & Al-Jasser, A. M. Infections Caused by *Stenotrophomonas maltophilia* in Recipients of Hematopoietic Stem Cell Transplantation. *Front Oncol* **4**, 232, doi:10.3389/fonc.2014.00232 (2014).

50 Chang, Y. T., Lin, C. Y., Chen, Y. H. & Hsueh, P. R. Update on infections caused by *Stenotrophomonas maltophilia* with particular attention to resistance mechanisms and therapeutic options. *Front Microbiol* **6**, 893, doi:10.3389/fmicb.2015.00893 (2015).

51 Zhanel, G. G. *et al.* Antimicrobial susceptibility of 22746 pathogens from Canadian hospitals: results of the CANWARD 2007-11 study. *J Antimicrob Chemother* **68 Suppl 1**, i7-22, doi:10.1093/jac/dkt022 (2013).

52 Walkty, A. *et al. In vitro* activity of plazomicin against 5,015 Gram-negative and Gram-positive clinical isolates obtained from patients in Canadian hospitals as part of the CANWARD study, 2011-2012. *Antimicrob Agents Chemother* **58**, 2554-2563, doi:10.1128/AAC.02744-13 (2014).

53 Naidu, P. & Smith, S. A review of 11 years of *Stenotrophomonas maltophilia* blood isolates at a tertiary care institute in Canada. *Canadian Journal of Infectious Diseases and Medical Microbiology* **23** (2012).

54 Denton, M. & Kerr, K. G. Microbiological and clinical aspects of infection associated with *Stenotrophomonas maltophilia*. *Clinical Microbiology Reviews* **11**, 57-80 (1998).

55 Velázquez-Acosta, C., Zarco-Márquez, S., Jiménez-Andrade, M. C., Volkow-Fernández, P. & Cornejo-Juárez, P. *Stenotrophomonas maltophilia* bacteremia and pneumonia at a tertiary-care oncology center: a review of 16 years. *Supportive Care in Cancer* **26**, 1953-1960, doi:10.1007/s00520-017-4032-x (2018).

56 Schwab, F., Geffers, C., Behnke, M. & Gastmeier, P. ICU mortality following ICU-acquired primary bloodstream infections according to the type of pathogen: A prospective cohort study in 937 Germany ICUs (2006-2015). *PLoS One* **13**, e0194210, doi:10.1371/journal.pone.0194210 (2018).

57 Kampmeier, S. *et al.* Evaluation of a *Stenotrophomonas maltophilia* bacteremia cluster in hematopoietic stem cell transplantation recipients using whole genome sequencing. *Antimicrobial Resistance and Infection Control* **6**, doi:10.1186/s13756-017-0276-y (2017).

58 Silbaq, F. S. Viable ultramicrocells in drinking water. *Journal of Applied Microbiology* **106**, 106-117, doi:doi:10.1111/j.1365-2672.2008.03981.x (2009).

59 Ahn, G. Y. *et al.* Pseudo-outbreak of *Stenotrophomonas maltophilia* due to contamination of bronchoscope. *The Korean journal of laboratory medicine* **27**, 205-209, doi:10.3343/kjlm.2007.27.3.205 (2007).

60 Wainwright, C. E. *et al.* Cough-generated aerosols of *Pseudomonas aeruginosa* and other Gram-negative bacteria from patients with cystic fibrosis. *Thorax* **64**, 926-931, doi:10.1136/thx.2008.112466 (2009).

61 Knibbs, L. D. *et al.* Viability of *Pseudomonas aeruginosa* in cough aerosols generated by persons with cystic fibrosis. *Thorax* **69**, 740-745, doi:10.1136/thoraxjnl-2014-205213 (2014).

62 Furlan, J. P. R. & Stehling, E. G. Detection of *blaPER* on an IncA/C plasmid in *Stenotrophomonas maltophilia* isolated from Brazilian soil. *Water, Air, & Soil Pollution* **229**, doi:10.1007/s11270-018-3799-9 (2018).

63 Kim, Y. J., Park, J. H. & Seo, K. H. Presence of *Stenotrophomonas maltophilia* exhibiting high genetic similarity to clinical isolates in final effluents of pig farm wastewater treatment plants. *Int J Hyg Environ Health* **221**, 300-307, doi:10.1016/j.ijheh.2017.12.002 (2018).

64 Arvanitidou, M., Vayona, A., Spanakis, N. & Tsakris, A. Occurrence and antimicrobial resistance of Gram-negative bacteria isolated in haemodialysis water and dialysate of renal units: results of a Greek multicentre study. *Journal of Applied Microbiology* **95**, 180-185, doi:10.1046/j.1365-2672.2003.01966.x (2003).

65 Berg, G., Eberl, L. & Hartmann, A. The rhizosphere as a reservoir for opportunistic human pathogenic bacteria. *Environmental Microbiology* **7**, 1673-1685, doi:10.1111/j.1462-2920.2005.00891.x (2005).

66 Qureshi, A., Mooney, L., Denton, M. & Kerr, K. G. *Stenotrophomonas maltophilia* in salad. *Emerg Infect Dis* **11**, 1157-1158, doi:10.3201/eid1107.040130 (2005).

67 Lin, L. *et al.* Antimicrobial resistance and genetic diversity in ceftazidime non-susceptible bacterial pathogens from ready-to-eat street foods in three Taiwanese cities. *Sci Rep* **7**, 15515, doi:10.1038/s41598-017-15627-8 (2017).

68 Falagas, M. E., Kastoris, A. C., Vouloumanou, E. K. & Dimopoulos, G. Community-acquired *Stenotrophomonas maltophilia* infections: a systematic review. *Eur J Clin Microbiol Infect Dis* **28**, 719-730, doi:10.1007/s10096-009-0709-5 (2009).

69 Chang, Y. T. *et al. Stenotrophomonas maltophilia* bloodstream infection: Comparison between community-onset and hospital-acquired infections. *Journal of Microbiology, Immunology and Infection* **47**, 28-35, doi:10.1016/j.jmii.2012.08.014 (2014).

70 Turrientes, M. C. *et al.* Polymorphic mutation frequencies of clinical and environmental *Stenotrophomonas maltophilia* populations. *Applied and Environmental Microbiology* **76**, 1746-1758, doi:10.1128/AEM.02817-09 (2010).

71 Garcia, C. A., Alcaraz, E. S., Franco, M. A. & Passerini de Rossi, B. N. Iron is a signal for *Stenotrophomonas maltophilia* biofilm formation, oxidative stress response, OMPs expression, and virulence. *Frontiers in Microbiology* **6** (2015).

72 An, S. Q. & Tang, J. L. The Ax21 protein influences virulence and biofilm formation in *Stenotrophomonas maltophilia*. *Arch Microbiol* **200**, 183-187, doi:10.1007/s00203-017-1433-7 (2018).

73 Crossman, L. C. *et al.* The complete genome, comparative and functional analysis of *Stenotrophomonas maltophilia* reveals an organism heavily shielded by drug resistance determinants. *Genome Biol* **9**, R74, doi:10.1186/gb-2008-9-4-r74 (2008).

74 Figueirêdo, P. M. S. *et al.* Cytotoxic activity of clinical *Stenotrophomonas maltophilia*. *Letters in Applied Microbiology* **43**, 443-449, doi:10.1111/j.1472-765X.2006.01965.x (2006).

75 Neal, D. J. & Wilkinson, S. G. Lipopolysaccharides from *Pseudomonas maltophilia* Structural Studies of the Side-Chain, Core, and Lipid-A Regions of the

Lipopolysaccharide from Strain NCTC 10257. *European Journal of Biochemistry* **128**, 143-149, doi:10.1111/j.1432-1033.1982.tb06944.x (1982).

76 Zhang, J. & Kong, F. Synthesis of a xylosylated rhamnose pentasaccharide: the repeating unit of the O-specific side chain of lipopolysaccharides from the reference strains for *Stenotrophomonas maltophilia* serogroup O18. *J. Carbohydr. Chem.* **21**, 89 –97 (2002).

77 McKay, G. A., Woods, D. E., MacDonald, K. L. & Poole, K. Role of phosphoglucomutase of *Stenotrophomonas maltophilia* in lipopolysaccharide biosynthesis, virulence, and antibiotic resistance. *Infect Immun* **71**, 3068-3075 (2003).

78 Huang, T. P., Somers, E. B. & Wong, A. C. L. Differential biofilm formation and motility associated with lipopolysaccharide/exopolysaccharide-coupled biosynthetic genes in *Stenotrophomonas maltophilia*. *Journal of Bacteriology* **188**, 3116-3120, doi:10.1128/JB.188.8.3116-3120.2006 (2006).

79 Brooke, J. S., Vo, A., Watts, P. & Davis, N. A. Mutation of a lipopolysaccharide synthesis gene results in increased biofilm of *Stenotrophomonas maltophilia* on plastic and glass surfaces. . *Ann. Microbiol.* **58**, 35– 40 (2008).

80 Vo, A., Davis, N. A., Watts, P. & Brooke, J. S. in *American Society for Microbiology.*

81 Pompilio, A. *et al.* Factors associated with adherence to and biofilm formation on polystyrene by *Stenotrophomonas maltophilia*: the role of cell surface hydrophobicity and motility. *FEMS Microbiol Lett* **287**, 41-47, doi:10.1111/j.1574-6968.2008.01292.x (2008).

82 Brooke, J. S. New strategies against *Stenotrophomonas maltophilia*: A serious worldwide intrinsically drug-resistant opportunistic pathogen. *Expert Review of Anti-Infective Therapy* **12**, 1-4, doi:10.1586/14787210.2014.864553 (2014).

83 Pompilio, A. *et al.* Adhesion to and biofilm formation on IB3-1 bronchial cells by *Stenotrophomonas maltophilia* isolates from cystic fibrosis patients. *BMC Microbiol* **10**, 102, doi:10.1186/1471-2180-10-102 (2010).

84 de Oliveira-Garcia, D. *et al.* Fimbriae and adherence of *Stenotrophomonas maltophilia* to epithelial cells and to abiotic surfaces. *Cellular Microbiology* **5**, 625-636, doi:10.1046/j.1462-5822.2003.00306.x (2003).

85 Lai, C. H. *et al.* Central venous catheter-related *Stenotrophomonas maltophilia* bacteraemia and associated relapsing bacteraemia in haematology and oncology patients. *Clin Microbiol Infect* **12**, 986-991, doi:10.1111/j.1469-0691.2006.01511.x (2006).

86 De Vidipo, L. A., De Marques, E. A., Puchelle, E. & Plotkowski, M. C. *Stenotrophomonas maltophilia* interaction with human epithelial respiratory cells *in vitro*. *Microbiol Immunol* **45**, 563-569 (2001).

87 O'Sullivan, B. P. & Freedman, S. D. Cystic fibrosis. *The Lancet* **373**, 1891-1904, doi:10.1016/S0140-6736(09)60327-5 (2009).

88 Cianciotto, N. P. & White, R. C. Expanding role of type II secretion in bacterial pathogenesis and beyond. *Infect Immun* **85**, e00014-00017, doi:10.1128/IAI.00014-17 (2017).

89 Tsirigotaki, A., De Geyter, J., Šoštarić, N., Economou, A. & Karamanou, S. Protein export though the bacterial Sec pathway. *Nature Reviews Microbiology* **15**, 21-36, doi:10.1038/nrmicro.2016.161 (2017).

90 Berks, B. C. in *Annual Review of Biochemistry* Vol. 84 843-864 (2015).

91 Costa, T. R. D. *et al.* Secretion systems in Gram-negative bacteria: Structural and mechanistic insights. *Nature Reviews Microbiology* **13**, 343-359, doi:10.1038/nrmicro3456 (2015).

92 Karaba, S. M., White, R. C. & Cianciotto, N. P. *Stenotrophomonas maltophilia* encodes a type II protein secretion system that promotes detrimental effects on lung epithelial cells. *Infect Immun* **81**, 3210-3219, doi:10.1128/IAI.00546-13 (2013).

93 DuMont, A. L., Karaba, S. M. & Cianciotto, N. P. Type II secretion-dependent degradative and cytotoxic activities mediated by *Stenotrophomonas maltophilia* serine

proteases StmPr1 and StmPr2. *Infect Immun* **83**, 3825-3837, doi:10.1128/IAI.00672-15 (2015).

94 DuMont, A. L. & Cianciotto, N. P. *Stenotrophomonas maltophilia* serine protease StmPr1 induces matrilysis, anoikis, and protease-activated receptor 2 activation in human lung epithelial cells. *Infect Immun* **85**, doi:10.1128/IAI.00544-17 (2017).

95 Ryan, R. P. *et al.* Interspecies signalling via the *Stenotrophomonas maltophilia* diffusible signal factor influences biofilm formation and polymyxin tolerance in *Pseudomonas aeruginosa*. *Mol Microbiol* **68**, 75-86, doi:10.1111/j.1365-2958.2008.06132.x (2008).

96 An, S. Q. & Tang, J. L. Diffusible signal factor signaling regulates multiple functions in the opportunistic pathogen *Stenotrophomonas maltophilia*. *BMC Res Notes* **11**, 569, doi:10.1186/s13104-018-3690-1 (2018).

97 Mett, H., Rosta, S., Schacher, B. & Frei, R. Outer membrane permeability and beta-lactamase content in *Pseudomonas maltophilia* clinical isolates and laboratory mutants. *Rev Infect Dis* **10**, 765-769 (1988).

98 Sánchez, M. B. & Martínez, J. L. SmQnr contributes to intrinsic resistance to quinolones in *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **54**, 580-581, doi:10.1128/AAC.00496-09 (2010).

99 Sánchez, M. B., Hernández, A., Rodríguez-Martínez, J. M., Martínez-Martínez, L. & Martínez, J. L. Predictive analysis of transmissible quinolone resistance indicates *Stenotrophomonas maltophilia* as a potential source of a novel family of Qnr determinants. *BMC Microbiology* **8**, doi:10.1186/1471-2180-8-148 (2008).

100 Chang, Y. C., Tsai, M. J., Huang, Y. W., Chung, T. C. & Yang, T. C. SmQnrR, a DeoR-type transcriptional regulator, negatively regulates the expression of Smqnr and SmtcrA in *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **66**, 1024-1028, doi:10.1093/jac/dkr049 (2011).

101 Okazaki, A. & Avison, M. B. Induction of L1 and L2 β-lactamase production in *Stenotrophomonas maltophilia* is dependent on an AmpR-type regulator. *Antimicrobial Agents and Chemotherapy* **52**, 1525-1528, doi:10.1128/AAC.01485-07 (2008).

102 Crowder, M. W., Walsh, T. R., Banovic, L., Pettit, M. & Spencer, J. Overexpression, purification, and characterization of the cloned metallo-β-lactamase L1 from *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **42**, 921-926 (1998).

103 Walsh, T. R., MacGowan, A. P. & Bennett, P. M. Sequence analysis and enzyme kinetics of the L2 serine β-lactamase from *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **41**, 1460-1464 (1997).

104 Okazaki, A. & Avison, M. B. Aph(3′)-IIc, an aminoglycoside resistance determinant from *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **51**, 359-360, doi:10.1128/AAC.00795-06 (2007).

105 Vetting, M. W. *et al.* Mechanistic and structural analysis of aminoglycoside N-acetyltransferase AAC(6')-Ib and its bifunctional, fluoroquinolone-active AAC(6')-Ib-cr variant. *Biochemistry* **47**, 9825-9835, doi:10.1021/bi800664x (2008).

106 Wang, C. H. *et al.* Comparisons between patients with trimethoprim-sulfamethoxazole-susceptible and trimethoprim-sulfamethoxazole-resistant *Stenotrophomonas maltophilia* monomicrobial bacteremia: A 10-year retrospective study. *J Microbiol Immunol Infect* **49**, 378-386, doi:10.1016/j.jmii.2014.06.005 (2016).

107 Chung, H. S. *et al.* Antimicrobial susceptibility of *Stenotrophomonas maltophilia* isolates from Korea, and the activity of antimicrobial combinations against the isolates. *Journal of Korean Medical Science* **28**, 62-66, doi:10.3346/jkms.2013.28.1.62 (2013).

108 Rhee, J. Y., Song, J. H. & Ko, K. S. Current situation of antimicrobial resistance and genetic differences in *Stenotrophomonas maltophilia* complex isolates by multilocus variable number of tandem repeat analysis. *Infect Chemother* **48**, 285-293, doi:10.3947/ic.2016.48.4.285 (2016).

109 Avison, M. B., von Heldreich, C. J., Higgins, C. S., Bennett, P. M. & Walsh, T. R. A TEM-2 beta-lactamase encoded on an active Tn1-like transposon in the genome of a clinical isolate of *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **46**, 879–884 (2000).

110 Avison, M. B., Higgins, C. S., von Heldreich, C. J., Bennett, P. M. & Walsh, T. R. Plasmid location and molecular heterogeneity of the L1 and L2 beta-lactamase genes of *Stenotrophomonas maltophilia*. *Antimicrob Agents Chemother* **45**, 413-419, doi:10.1128/AAC.45.2.413-419.2001 (2001).

111 Barbolla, R. *et al.* Class 1 integrons increase trimethoprim-sulfamethoxazole MICs against epidemiologically unrelated *Stenotrophomonas maltophilia* isolates. *Antimicrobial Agents and Chemotherapy* **48**, 666-669, doi:10.1128/AAC.48.2.666-669.2004 (2004).

112 Chung, H. S. *et al.* The *sul1* gene in *Stenotrophomonas maltophilia* with high-level resistance to trimethoprim/sulfamethoxazole. *Annals of Laboratory Medicine* **35**, 246-249, doi:10.3343/alm.2015.35.2.246 (2015).

113 Toleman, M. A., Bennett, P. M., Bennett, D. M., Jones, R. N. & Walsh, T. R. Global emergence of trimethoprim/sulfamethoxazole resistance in *Stenotrophomonas maltophilia* mediated by acquisition of *sul* genes. *Emerg Infect Dis* **13**, 559-565, doi:10.3201/eid1304.061378 (2007).

114 Hu, L. F. *et al. Stenotrophomonas maltophilia* resistance to trimethoprim/sulfamethoxazole mediated by acquisition of *sul* and *dfrA* genes in a plasmid-mediated class 1 integron. *International Journal of Antimicrobial Agents* **37**, 230-234, doi:10.1016/j.ijantimicag.2010.10.025 (2011).

115 Huang, Y. W., Hu, R. M. & Yang, T. C. Role of the pcm-tolCsm operon in the multidrug resistance of *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **68**, 1987-1993, doi:10.1093/jac/dkt148 (2013).

116 Lin, Y. T., Huang, Y. W., Chen, S. J., Chang, C. W. & Yang, T. C. The SmeYZ efflux pump of *Stenotrophomonas maltophilia* contributes to drug resistance, virulence-related characteristics, and virulence in mice. *Antimicrob Agents Chemother* **59**, 4067-4073, doi:10.1128/AAC.00372-15 (2015).

117 Sanchez, M. B. Antibiotic resistance in the opportunistic pathogen *Stenotrophomonas maltophilia*. *Front Microbiol* **6**, 658, doi:10.3389/fmicb.2015.00658 (2015).

118 Putman, M., van Veen, H. W. & Konings, W. N. Molecular properties of bacterial multidrug transporters. *Microbiol Mol Biol Rev* **64**, 672-693, doi:10.1128/MMBR.64.4.672-693.2000 (2000).

119 Alonso, A. & Martinez, J. L. Cloning and characterization of SmeDEF, a novel multidrug efflux pump from *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **44**, 3079-3086, doi:10.1128/AAC.44.11.3079-3086.2000 (2000).

120 Chen, C. H. *et al.* Contribution of resistance-nodulation-division efflux pump operon smeU1-V-W-U2-X to multidrug resistance of *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **55**, 5826-5833, doi:10.1128/AAC.00317-11 (2011).

121 García-León, G. *et al.* A function of SmeDEF, the major quinolone resistance determinant of *Stenotrophomonas maltophilia*, is the colonization of plant roots. *Applied and Environmental Microbiology* **80**, 4559-4565, doi:10.1128/AEM.01058-14 (2014).

122 García-León, G. *et al.* High-level quinolone resistance is associated with the overexpression of smeVWX in *Stenotrophomonas maltophilia* clinical isolates. *Clinical Microbiology and Infection* **21**, 464-467, doi:10.1016/j.cmi.2015.01.007 (2015).

123 Gould, V. C., Okazaki, A. & Avison, M. B. Coordinate hyperproduction of SmeZ and SmeJK efflux pumps extends drug resistance in *Stenotrophomonas maltophilia*. *Antimicrobial Agents and Chemotherapy* **57**, 655-657, doi:10.1128/AAC.01020-12 (2013).

124 Wu, C. J. *et al.* Role of smeU1VWU2X operon in alleviation of oxidative stresses and occurrence of sulfamethoxazole-trimethoprim-resistant mutants in *Stenotrophomonas maltophilia*. *Antimicrob Agents Chemother* **62**, doi:10.1128/AAC.02114-17 (2018).

125 Lin, Y. T., Huang, Y. W., Liou, R. S., Chang, Y. C. & Yang, T. C. MacABCsm, an ABC-type tripartite efflux pump of *Stenotrophomonas maltophilia* involved in drug resistance, oxidative and envelope stress tolerances and biofilm formation. *J Antimicrob Chemother* **69**, 3221-3226, doi:10.1093/jac/dku317 (2014).

126 Huang, Y. W., Hu, R. M., Chu, F. Y., Lin, H. R. & Yang, T. C. Characterization of a major facilitator superfamily (MFS) tripartite efflux pump emrcabsm from *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **68**, 2498-2505, doi:10.1093/jac/dkt250 (2013).

127 Hu, R. M., Liao, S. T., Huang, C. C., Huang, Y. W. & Yang, T. C. An inducible fusaric acid tripartite efflux pump contributes to the fusaric acid resistance in *Stenotrophomonas maltophilia*. *PLoS ONE* **7**, doi:10.1371/journal.pone.0051053 (2012).

128 Sanchez, M. B., Hernandez, A. & Martinez, J. L. *Stenotrophomonas maltophilia* drug resistance. *Future Microbiol* **4**, 655-660, doi:10.2217/fmb.09.45 (2009).

129 Tada, T. *et al.* Identification of a novel 6'-N-aminoglycoside acetyltransferase, AAC(6')-Iak, from a multidrug-resistant clinical isolate of *Stenotrophomonas maltophilia*. *Antimicrob Agents Chemother* **58**, 6324-6327, doi:10.1128/AAC.03354-14 (2014).

130 Li, X. Z., Zhang, L., McKay, G. A. & Poole, K. Role of the acetyltransferase AAC(6′)-Iz modifying enzyme in aminoglycoside resistance in *Stenotrophomonas maltophilia*. *Journal of Antimicrobial Chemotherapy* **51**, 803-811, doi:10.1093/jac/dkg148 (2003).

131 Golkar, Z., Bagasra, O. & Pace, D. G. Bacteriophage therapy: a potential solution for the antibiotic resistance crisis. *J Infect Dev Ctries* **8**, 129-136, doi:10.3855/jidc.3573 (2014).

132 Burrowes, B., Harper, D. R., Anderson, J., McConville, M. & Enright, M. C. Bacteriophage therapy: potential uses in the control of antibiotic-resistant pathogens. *Expert Review of Anti-Infective Therapy* **9**, 775-785, doi:10.1586/eri.11.90 (2011).

133 Cooper, C. J., Khan Mirzaei, M. & Nilsson, A. S. Adapting drug approval pathways for bacteriophage-based therapeutics. *Front Microbiol* **7**, 1209, doi:10.3389/fmicb.2016.01209 (2016).

134 Sulakvelidze, A., Alavidze, Z. & Morris, J. G., Jr. Bacteriophage therapy. *Antimicrob Agents Chemother* **45**, 649-659, doi:10.1128/AAC.45.3.649-659.2001 (2001).

135 Shan, J. *et al.* Bacteriophages are more virulent to bacteria with human cells than they are in bacterial culture; insights from HT-29 cells. *Sci Rep* **8**, 5091, doi:10.1038/s41598-018-23418-y (2018).

136 Meader, E., Mayer, M. J., Steverding, D., Carding, S. R. & Narbad, A. Evaluation of bacteriophage therapy to control *Clostridium difficile* and toxin production in an *in vitro* human colon model system. *Anaerobe* **22**, 25-30, doi:10.1016/j.anaerobe.2013.05.001 (2013).

137 Yu, L. *et al.* A guard-killer phage cocktail effectively lyses the host and inhibits the development of phage-resistant strains of *Escherichia coli*. *Appl Microbiol Biotechnol* **102**, 971-983, doi:10.1007/s00253-017-8591-z (2018).

138 Singla, S., Harjai, K., Katare, O. P. & Chhibber, S. Encapsulation of bacteriophage in liposome accentuates its entry in to macrophage and shields it from neutralizing antibodies. *PLOS ONE* **11**, e0153777, doi:10.1371/journal.pone.0153777 (2016).

139 Chanishvili, N. in *Advances in Virus Research, Vol 83: Bacteriophages, Pt B* Vol. 83 *Advances in Virus Research* (eds M. Lobocka & W. T. Szybalski) 3-40 (2012).

140 Merabishvili, M. *et al.* Quality-controlled small-scale production of a well-defined bacteriophage cocktail for use in human clinical trials. *Plos One* **4**, doi:10.1371/journal.pone.0004944 (2009).

141 Rhoads, D. D. *et al.* Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. *Journal of wound care* **18**, 237-238, 240-233 (2009).

142 Wright, A., Hawkins, C. H., Anggard, E. E. & Harper, D. R. A controlled clinical trial of a therapeutic bacteriophage preparation in chonic otitis due to antibiotic-resistant *Pseudomonas aeruginosa*; a preliminary report of efficacy. *Clinical Otolaryngology* **34**, 349-357, doi:10.1111/j.1749-4486.2009.01973.x (2009).

143 Saussereau, E. *et al.* Effectiveness of bacteriophages in the sputum of cystic fibrosis patients. *Clin Microbiol Infect* **20**, O983-990, doi:10.1111/1469-0691.12712 (2014).

144 Sarker, S. A. *et al.* Oral phage therapy of acute bacterial diarrhea with two coliphage preparations: a randomized trial in children from bangladesh. *EBioMedicine* **4**, 124-137, doi:10.1016/j.ebiom.2015.12.023 (2016).

145 Fish, R. *et al.* Bacteriophage treatment of intransigent diabetic toe ulcers: A case series. *Journal of Wound Care* **25**, S27-S33, doi:10.12968/jowc.2016.25.7.S27 (2016).

146 Leitner, L. *et al.* Bacteriophages for treating urinary tract infections in patients undergoing transurethal resection of the prostate: a randomized, placebo-controlled, double-blind clinical trial. *BMC Urol* **17**, 90, doi:10.1186/s12894-017-0283-6 (2017).

147 Waldor, M. K. & Mekalanos, J. J. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910-1914 (1996).

148 Brussow, H. & Hendrix, R. Phage genomics: Small is beautiful. *Cell* **108**, 13-16 (2002).

149 Hayashi, T. *et al.* Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**, 11-22 (2001).

150 Arndt, D. *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**, W16-21, doi:10.1093/nar/gkw387 (2016).

151 Brussow, H., Canchaya, C. & Hardt, W. D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**, 560-602, doi:10.1128/MMBR.68.3.560-602.2004 (2004).

152 Moillo, A. M. Isolation of a transducing phage forming plaques on *Pseudomonas maltophilia* and *Pseudomonas aeruginosa*. *Genet Res* **21**, 287-289, doi:10.1017/s0016672300013471 (1973).

153 Chang, H. C. *et al.* Isolation and characterization of novel giant *Stenotrophomonas maltophilia* phage phi SMA5. *Applied and Environmental Microbiology* **71**, 1387-1393, doi:10.1128/aem.71.3.1387-1393.2005 (2005).

154 Hagemann, M., Hasse, D. & Berg, G. Detection of a phage genome carrying a zonula occludens like toxin gene (*zot*) in clinical isolates of *Stenotrophomonas maltophilia*. *Arch Microbiol* **185**, 449-458, doi:10.1007/s00203-006-0115-7 (2006).

155 Chen, C.-R. *et al.* Characterization of a novel T4-type *Stenotrophomonas maltophilia* virulent phage Smp14. *Archives of Microbiology* **188**, 191-197, doi:10.1007/s00203-007-0238-5 (2007).

156 Garcia, P. *et al.* Isolation of new *Stenotrophomonas* bacteriophages and genomic characterization of temperate phage S1. *Appl Environ Microbiol* **74**, 7552-7560, doi:10.1128/AEM.01709-08 (2008).

157 Fan, H. *et al.* Complete genome sequence of IME13, a *Stenotrophomonas maltophilia* bacteriophage with large burst size and unique plaque polymorphism. *Journal of Virology* **86**, 11392-11393, doi:10.1128/jvi.01908-12 (2012).

158 Huang, Y. *et al.* Complete genome sequence of IME15, the first T7-like bacteriophage lytic to pan-antibiotic- resistant *Stenotrophomonas maltophilia*. *Journal of Virology* **86**, 13839-13840, doi:10.1128/jvi.02661-12 (2012).

159 Liu, J., Liu, Q., Shen, P. & Huang, Y. P. Isolation and characterization of a novel filamentous phage from *Stenotrophomonas maltophilia*. *Arch Virol* **157**, 1643-1650, doi:10.1007/s00705-012-1305-z (2012).

160 Zhang, J. & Li, X. [Biological characteristics of phage SM1 for *Stenotrophomonas maltophilia* and its effect in animal infection model]. *Zhejiang Da Xue Xue Bao Yi Xue Ban* **42**, 331-336 (2013).

161 Petrova, M., Shcherbatova, N., Kurakov, A. & Mindlin, S. Genomic characterization and integrative properties of phiSMA6 and phiSMA7, two novel filamentous bacteriophages of *Stenotrophomonas maltophilia*. *Archives of Virology* **159**, 1293-1303, doi:10.1007/s00705-013-1882-5 (2014).

162 Lee, C.-N., Tseng, T.-T., Chang, H.-C., Lin, J.-W. & Weng, S.-F. Genomic sequence of temperate phage Smp131 of *Stenotrophomonas maltophilia* that has similar prophages in xanthomonads. *Bmc Microbiology* **14**, doi:10.1186/1471-2180-14-17 (2014).

163 Peters, D. L., Lynch, K. H., Stothard, P. & Dennis, J. J. The isolation and characterization of two *Stenotrophomonas maltophilia* bacteriophages capable of cross-taxonomic order infectivity. *BMC Genomics* **16**, 664, doi:10.1186/s12864-015-1848-y (2015).

164 Peters, D. L., Stothard, P. & Dennis, J. J. The isolation and characterization of *Stenotrophomonas maltophilia* T4-like bacteriophage DLP6. *PLoS One* **12**, e0173341, doi:10.1371/journal.pone.0173341 (2017).

165 Bondy-Denomy, J. *et al.* Prophages mediate defense against phage infection though diverse mechanisms. *ISME J* **10**, 2854-2866, doi:10.1038/ismej.2016.79 (2016).

166 Peters, D. L. & Dennis, J. J. Complete genome sequence of temperate *Stenotrophomonas maltophilia* bacteriophage DLP5. *Genome Announc* **6**, doi:10.1128/genomeA.00073-18 (2018).

167 Seed, K. D. & Dennis, J. J. Isolation and characterization of bacteriophages of the *Burkholderia cepacia* complex. *FEMS Microbiol Lett* **251**, 273-280, doi:10.1016/j.femsle.2005.08.011 (2005).

168 Corporation, P. DNA isolation from Lambda lysates using the Wizard® DNA clean-up system. (2006).

169 Lynch, K. H., Abdu, A. H., Schobert, M. & Dennis, J. J. Genomic characterization of JG068, a novel virulent podovirus active against *Burkholderia cenocepacia*. *BMC Genomics* **14**, 574, doi:10.1186/1471-2164-14-574 (2013).

170 Lynch, K. H., Seed, K. D., Stothard, P. & Dennis, J. J. Inactivation of *Burkholderia cepacia* complex phage KS9 gp41 identifies the phage repressor and generates lytic virions. *J Virol* **84**, 1276-1288, doi:10.1128/JVI.01843-09 (2010).

171 Lynch, K. H., Stothard, P. & Dennis, J. J. Genomic analysis and relatedness of P2-like phages of the *Burkholderia cepacia* complex. *Bmc Genomics* **11**, doi:10.1186/1471-2164-11-599 (2010).

172 Delcher, A. L., Bratke, K. A., Powers, E. C. & Salzberg, S. L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**, 673-679, doi:10.1093/bioinformatics/btm009 (2007).

173 Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649, doi:10.1093/bioinformatics/bts199 (2012).

174 Besemer, J., Lomsadze, A. & Borodovsky, M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* **29**, 2607-2618 (2001).

175 Marchler-Bauer, A. *et al.* CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res* **39**, D225-229, doi:10.1093/nar/gkq1189 (2011).

176 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

177 Krzywinski, M. *et al.* Circos: An information aesthetic for comparative genomics. *Genome Research* **19**, 1639-1645, doi:10.1101/gr.092759.109 (2009).

178 Delcher, A. L., Phillippy, A., Carlton, J. & Salzberg, S. L. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Research* **30**, 2478-2483, doi:10.1093/nar/30.11.2478 (2002).

179 Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-580, doi:10.1006/jmbi.2000.4315 (2001).

180 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high thoughput. *Nucleic Acids Res* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

181 Lynch, K. H., Stothard, P. & Dennis, J. J. Comparative analysis of two phenotypically-similar but genomically-distinct *Burkholderia* ceno*cepacia*-specific bacteriophages. *BMC Genomics* **13**, 223, doi:10.1186/1471-2164-13-223 (2012).

182 Sullivan, M. B., Waterbury, J. B. & Chisholm, S. W. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424**, 1047-1051, doi:10.1038/nature01929 (2003).

183 Watkins, S. C., Smith, J. R., Hayes, P. K. & Watts, J. E. M. Characterisation of host growth after infection with a broad-range freshwater cyanopodophage. *Plos One* **9**, doi:10.1371/journal.pone.0087339 (2014).

184 Weitz, J. S. *et al.* Phage-bacteria infection networks. *Trends in Microbiology* **21**, 82-91, doi:10.1016/j.tim.2012.11.003 (2013).

185 Karumidze, N. *et al.* Characterization of lytic *Pseudomonas aeruginosa* bacteriophages via biological properties and genomic sequences. *Applied Microbiology and Biotechnology* **94**, 1609-1617, doi:10.1007/s00253-012-4119-8 (2012).

186 Kwan, T., Liu, J., DuBow, M., Gros, P. & Pelletier, J. Comparative genomic analysis of 18 *Pseudomonas aeruginosa* bacteriophages. *Journal of Bacteriology* **188**, 1184-1187, doi:10.1128/jb.188.3.1184-1187.2006 (2006).

187 Pellegrin, V., Juretschko, S., Wagner, M. & Cottenceau, G. Morphological and biochemical properties of a *Sphaerotilus* sp. isolated from paper mill slimes. *Applied and Environmental Microbiology* **65**, 156-162 (1999).

188 Berry, J., Summer, E. J., Struck, D. K. & Young, R. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. *Molecular Microbiology* **70**, 341-351, doi:10.1111/j.1365-2958.2008.06408.x (2008).

189 Berry, J., Savva, C., Holzenburg, A. & Young, R. The lambda spanin components Rz and Rz1 undergo tertiary and quaternary rearrangements upon complex formation. *Protein Science* **19**, 1967-1977, doi:10.1002/pro.485 (2010).

190 Summer, E. J. *et al.* Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. *J Mol Biol* **373**, 1098-1112, doi:10.1016/j.jmb.2007.08.045 (2007).

191 Juncker, A. S. *et al.* Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci* **12**, 1652-1662, doi:10.1110/ps.0303703 (2003).

192 Soding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* **33**, W244-W248, doi:10.1093/nar/gki408 (2005).

193 Owttrim, G. W. RNA helicases: Diverse roles in prokaryotic response to abiotic stress. *Rna Biology* **10**, 96-110, doi:10.4161/rna.22638 (2013).

194 Mechold, U., Potrykus, K., Murphy, H., Murakami, K. S. & Cashel, M. Differential regulation by ppGpp versus pppGpp in *Escherichia coli*. *Nucleic Acids Research* **41**, 6175-6189, doi:10.1093/nar/gkt302 (2013).

195 Magnusson, L. U., Farewell, A. & Nystrom, T. ppGpp: a global regulator in *Escherichia coli*. *Trends in Microbiology* **13**, 236-242, doi:10.1016/j.tim.2005.03.008 (2005).

196 Maciag, M., Kochanowska, M., Lyzen, R., Wegrzyn, G. & Szalewska-Palasz, A. ppGpp inhibits the activity of *Escherichia coli* DnaG primase. *Plasmid* **63**, 61-67, doi:10.1016/j.plasmid.2009.11.002 (2010).

197 Gross, M., Marianovsky, I. & Glaser, G. MazG - a regulator of programmed cell death in *Escherichia coli*. *Molecular Microbiology* **59**, 590-601, doi:10.1111/j.1365-2958.2005.04956.x (2006).

198 Bryan, M. J. *et al.* Evidence for the intense exchange of mazg in marine cyanophages by horizontal gene transfer. *Plos One* **3**, doi:10.1371/journal.pone.0002048 (2008).

199 Clokie, M. R. J. & Mann, N. H. Marine cyanophages and light. *Environmental Microbiology* **8**, 2074-2082, doi:10.1111/j.1462-2920.2006.01171.x (2006).

200 Lewenza, S. *et al.* Construction of a mini-Tn5-luxCDABE mutant library in *Pseudomonas aeruginosa* PAO1: a tool for identifying differentially regulated genes. *Genome Res* **15**, 583-589, doi:10.1101/gr.3513905 (2005).

201 Roessner, C. A. & Ihler, G. M. Proteinase sensitivity of bacteriophage lambda tail proteins gpJ and pH in complexes with the lambda receptor. *J Bacteriol* **157**, 165-170 (1984).

202 Rees, C. & Botsaris, G. in *Understanding Tuberculosis - Global Experiences and Innovative Approaches to the Diagnosis* Ch. 6, (2012).

203 Pedulla, M. L. *et al.* Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**, 171-182 (2003).

204 Spinelli, S., Veesler, D., Bebeacua, C. & Cambillau, C. Structures and host-adhesion mechanisms of lactococcal siphophages. *Front Microbiol* **5**, 3, doi:10.3389/fmicb.2014.00003 (2014).

205 Dieterle, M. E., Spinelli, S., Sadovskaya, I., Piuri, M. & Cambillau, C. Evolved distal tail carbohydrate binding modules of *Lactobacillus* phage J-1: a novel type of anti-receptor widespread among lactic acid bacteria phages. *Mol Microbiol* **104**, 608-620, doi:10.1111/mmi.13649 (2017).

206 Dieterle, M. E. *et al.* Exposing the secrets of two well-known *Lactobacillus casei* phages, J-1 and PL-1, by genomic and structural analysis. *Appl Environ Microbiol* **80**, 7107-7121, doi:10.1128/AEM.02771-14 (2014).

207 Dieterle, M. E., Jacobs-Sera, D., Russell, D., Hatfull, G. & Piuri, M. Complete genome sequences of *Lactobacillus* phages J-1 and PL-1. *Genome Announc* **2**, doi:10.1128/genomeA.00998-13 (2014).

208 Matilla, M. A. & Salmond, G. P. C. Bacteriophage phi MAM1, a *Viunalikevirus*, is a broad-host-range, high-efficiency generalized transducer that infects environmental and clinical isolates of the enterobacterial genera *Serratia* and *Kluyvera*. *Applied and Environmental Microbiology* **80**, 6446-6457, doi:10.1128/aem.01546-14 (2014).

209 Bielke, L., Higgins, S., Donoghue, A., Donoghue, D. & Hargis, B. M. *Salmonella* host range of bacteriophages that infect muftiple genera. *Poultry Science* **86**, 2536-2540, doi:10.3382/ps.2007-00250 (2007).

210 Park, M. *et al.* Characterization and comparative genomic analysis of a novel bacteriophage, SFP10, simultaneously inhibiting both *Salmonella enterica* and *Escherichia coli* O157:H7. *Applied and Environmental Microbiology* **78**, 58-69, doi:10.1128/aem.06231-11 (2012).

211 Kim, M. & Ryu, S. Characterization of a T5-Like coliphage, SPC35, and differential development of resistance to SPC35 in *Salmonella enterica* serovar *Typhimurium* and *Escherichia coli*. *Applied and Environmental Microbiology* **77**, 2042-2050, doi:10.1128/aem.02504-10 (2011).

212 Kutter, E. *et al.* Phage therapy in clinical practice: treatment of human infections. *Current Pharmaceutical Biotechnology* **11**, 69-86, doi:10.2174/138920110790725401 (2010).

213 Seed, K. D. & Dennis, J. J. Experimental bacteriophage therapy increases survival of *Galleria mellonella* larvae infected with clinically relevant strains of the *Burkholderia cepacia* complex. *Antimicrobial Agents and Chemotherapy* **53**, 2205-2208, doi:10.1128/aac.01166-08 (2009).

214 Ryan, E. M., Alkawareek, M. Y., Donnelly, R. F. & Gilmore, B. F. Synergistic phage-antibiotic combinations for the control of *Escherichia coli* biofilms *in vitro*. *Fems*

Immunology and Medical Microbiology **65**, 395-398, doi:10.1111/j.1574-695X.2012.00977.x (2012).

215 Bankevich, A. *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**, 455-477, doi:10.1089/cmb.2012.0021 (2012).

216 Boulanger, P. in *Bacteriophages: Methods and Protocols, Volume 2 Molecular and Applied Aspects* (eds Martha R. J. Clokie & Andrew M. Kropinski) 227-238 (Humana Press, 2009).

217 Drummond, A. e. a. Geneious. (2013).

218 Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**, 845-858, doi:10.1038/nprot.2015.053 (2015).

219 Remmert, M., Biegert, A., Hauser, A. & Soding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods* **9**, 173-175, doi:10.1038/nmeth.1818 (2011).

220 Zimmermann, L. *et al.* A completely reimplemented mpi bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol.* **17**, 30587-30589 (2018).

221 Yang, J. & Zhang, Y. Protein structure and function prediction using I-TASSER. *Curr Protoc Bioinformatics* **52**, 5 8 1-15, doi:10.1002/0471250953.bi0508s52 (2015).

222 Schattner, P., Brooks, A. N. & Lowe, T. M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research* **33**, W686-W689, doi:10.1093/nar/gki366 (2005).

223 Ackermann, H. W. & Eisenstark, A. The present state of phage taxonomy. *Intervirology* **3**, 201-219, doi:10.1159/000149758 (1974).

224 Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Res* **42**, D222-230, doi:10.1093/nar/gkt1223 (2014).

225 Fan, N., Cutting, S. & Losick, R. Characterization of the *Bacillus subtilis* sporulation gene *spoVK*. *J Bacteriol* **174**, 1053-1054 (1992).

226 Whitmore, S. E. & Lamont, R. J. Tyrosine phosphorylation and bacterial virulence. *Int J Oral Sci* **4**, 1-6, doi:10.1038/ijos.2012.6 (2012).

227 Bair, C. & Black, L. W. Exclusion of glucosyl-hydroxymethylcytosine DNA containing bacteriophages. *J Mol Biol* **366**, 779–789 (2007).

228 Allison, G. E. & Verma, N. K. Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* **8**, 17-23 (2000).

229 Pradel, E., Parker, C. T. & Schnaitman, C. A. Structures of the *rsfaB*, *rfaI*, *rfaJ*, and *rfaS* genes of *Escherichia coli* K-12 and their roles in assembly of the lipopolysaccharide core. *Journal of Bacteriology* **174**, 4736-4745 (1992).

230 *Viral Molecular Machines*. (Springer, 2012).

231 Lokareddy, R. K. *et al.* Portal protein functions akin to a DNA-sensor that couples genome-packaging to icosahedral capsid maturation. *Nat Commun* **8**, 14310, doi:10.1038/ncomms14310 (2017).

232 Miller, J. L. *et al.* Selective ligand recognition by a diversity-generating retroelement variable protein. *PLoS Biol* **6**, e131, doi:10.1371/journal.pbio.0060131 (2008).

233 Miller, J. L. *Understanding receptor specificity though the massively variable major tropism determinant of Bordetella bacteriophage* Ph.D. thesis, University of California, (2006).

234 Belcaid, M., Bergeron, A. & Poisson, G. The evolution of the tape measure protein: units, duplications and losses. *BMC Bioinformatics* **12 Suppl 9**, S10, doi:10.1186/1471-2105-12-S9-S10 (2011).

235 Struck, A. W., Thompson, M. L., Wong, L. S. & Micklefield, J. S-adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and

other biotechnological applications. *Chembiochem* **13**, 2642-2655, doi:10.1002/cbic.201200556 (2012).

236 Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chonicle of convergence. *Trends in Biochemical Sciences* **28**, 329-335, doi:10.1016/s0968-0004(03)00090-2 (2003).

237 Wipf, J. R. K., Schwendener, S., Nielsen, J. B., Westh, H. & Perreten, V. The new macrolide-lincosamide-streptogramin b resistance gene *erm*(45) is located within a genomic island in *Staphylococcus fleurettii*. *Antimicrobial Agents and Chemotherapy* **59**, 3578-3581, doi:10.1128/aac.00369-15 (2015).

238 Prunier, A. L., Malbruny, B., Tande, D., Picard, B. & Leclercq, R. Clinical isolates of *Staphylococcus aureus* with ribosomal mutations conferring resistance to macrolides. *Antimicrobial Agents and Chemotherapy* **46**, 3054-3056, doi:10.1128/aac.46.9.3054-3056.2002 (2002).

239 Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008, doi:10.1038/msb4100050 (2006).

240 Hanahan, D., Jessee, J. & Bloom, F. R. Plasmid transformation of *Escherichia coli* and other bacteria. *Methods Enzymol* **204**, 63-113 (1991).

241 Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M. & Peterson, K. M. pBBR1MCS: A broad-host-range cloning vector. *Biotechniques* **16**, 800–802 (1994).

242 Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J. & Wishart, D. S. PHAST: a fast phage search tool. *Nucleic Acids Res* **39**, W347-352, doi:10.1093/nar/gkr485 (2011).

243 Wiegand, I., Hilpert, K. & Hancock, R. E. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc* **3**, 163-175, doi:10.1038/nprot.2007.521 (2008).

244 Andrade-Dominguez, A. & Kolter, R. Complete genome sequence of *Pseudomonas aeruginosa* phage AAT-1. *Genome Announc* **4**, doi:10.1128/genomeA.00165-16 (2016).

245 Sepulveda-Robles, O., Kameyama, L. & Guarneros, G. High diversity and novel species of *Pseudomonas aeruginosa* bacteriophages. *Appl Environ Microbiol* **78**, 4510-4515, doi:10.1128/AEM.00065-12 (2012).

246 Iyer, L. M., Babu, M. M. & Aravind, L. The HIRAN domain and recruitment of chromatin remodeling and repair activities to damaged DNA. *Cell Cycle* **5**, 775-782, doi:10.4161/cc.5.7.2629 (2006).

247 Hooton, S. P. T. & Connerton, I. F. *Campylobacter jejuni* acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. *Frontiers in Microbiology* **5** (2015).

248 Tseng, M. J., He, P., Hilfinger, J. M. & Greenberg, G. R. Bacteriophage T4 *nrdA* and *nrdB* genes, encoding ribonucleotide reductase, are expressed both separately and coordinately: Characterization of the *nrdb* promoter. *Journal of Bacteriology* **172** (1990).

249 Mikoulinskaia, G. V. *et al.* Purification and characterization of the deoxynucleoside monophosphate kinase of bacteriophage T5. *Protein Expr Purif* **27**, 195-201 (2003).

250 Duckworth, D. H. & Bessman, M. J. The enzymology of virus-infected bacteria. X. A biochemical-genetic study of the deoxynucleotide kinase induced by wild type and amber mutants of phage T4. *J Biol Chem* **242**, 2877-2885 (1967).

251 Hardy, L. W. *et al.* Atomic structure of thymidylate synthase: target for rational drug design. *Science* **235**, 448-455 (1987).

252 Sarkar, S. K., Chowdhury, C. & Ghosh, A. S. Deletion of penicillin-binding protein 5 (PBP5) sensitises *Escherichia coli* cells to β-lactam agents. *Int. J. Antimicrob. Agents* **35**, 244–249 (2017).

253 Young, R. Phage lysis: three steps, three choices, one outcome. *J Microbiol* **52**, 243-258, doi:10.1007/s12275-014-4087-z (2014).

254 Severi, E., Hood, D. W. & Thomas, G. H. Sialic acid utilization by bacterial pathogens. *Microbiology* **153**, 2817-2822, doi:10.1099/mic.0.2007/009480-0 (2007).

255 Ahern, S. J., Das, M., Bhowmick, T. S., Young, R. & Gonzalez, C. F. Characterization of novel virulent broad-host-range phages of *Xylella fastidiosa* and *Xanthomonas*. *J Bacteriol* **196**, 459-471, doi:10.1128/JB.01080-13 (2014).

256 Dong, Z., Liu, J. & Peng, D. Accession No. KX241618.1 - *Xanthomonas* phage Xoo-sp2, complete genome. *National Library of Medicine (US), National Center for Biotechnology Information*.

257 Quin, M. B. *et al.* The bacterial stressosome: a modular system that has been adapted to control secondary messenger signaling. *Structure* **20**, 350-363, doi:10.1016/j.str.2012.01.003 (2012).

258 Inoue, T. *et al.* Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. *J Bacteriol* **189**, 950-957, doi:10.1128/JB.01294-06 (2007).

259 Frelin, O. *et al.* A directed-overflow and damage-control N-glycosidase in riboflavin biosynthesis. *Biochem J* **466**, 137-145, doi:10.1042/BJ20141237 (2015).

260 Savioz, A., Jeenes, D. J., Kocher, H. P. & Haas, D. Comparison of *proC* and other housekeeping genes of *Pseudomonas aeruginosa* with their counterparts in *Escherichia coli*. *Gene* **86,**, 107–111 (1990).

261 Harris, R. S. *Improved pairwise alignment of genomic DNA* Ph.D. thesis, Penn State University, (2007).

262 Schwartz, S. *et al.* Human-mouse alignments with BLASTZ. *Genome Res* **13**, 103-107, doi:10.1101/gr.809403 (2003).

263 Makino, Y. *et al.* Chromosome mapping and expression of human tip49 family genes. *DNA Sequence* **11**, 145-148, doi:10.3109/10425170009033982 (2009).

264 Meier-Dieter, U., Starman, R., Barr, K., Mayer, H. & Rick, P. D. Biosynthesis of enterobacterial common antigen in *Escherichia coli*. Biochemical characterization of Tn10 insertion mutants defective in enterobacterial common antigen synthesis. *J Biol Chem* **265**, 13490-13497 (1990).

265 Hoe, S. *et al.* Respirable bacteriophages for the treatment of bacterial lung infections. *Journal of Aerosol Medicine and Pulmonary Drug Delivery* **26**, 317-335, doi:10.1089/jamp.2012.1001 (2013).

266 Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**, doi:10.1186/1471-2105-11-119 (2010).

267 Aziz, R. K. *et al.* The RAST Server: Rapid annotations using subsystems technology. *Nucleic Acids Research* **9** (2008).

268 Overbeek, R. *et al.* The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Research* **42**, D206-D214, doi:10.1093/nar/gkt1226 (2014).

269 Brettin, T. *et al.* RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5**, 8365, doi:10.1038/srep08365 (2015).

270 Gautheret, D. & Lambert, A. Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles. *Journal of Molecular Biology* **313**, 1003-1011, doi:10.1006/jmbi.2001.5102 (2001).

271 Macke, T. J. *et al.* RNAMotif, an RNA secondary structure definition and search algorithm. *Nucleic Acids Research* **29**, 4724-4735, doi:10.1093/nar/29.22.4724 (2001).

272 Lesnik, E. A. *et al.* Prediction of rho-independent transcriptional terminators in *Escherichia coli*. *Nucleic Acids Research* **29**, 3583-3594, doi:10.1093/nar/29.17.3583 (2001).

273 Lavigne, R., Sun, W. D. & Volckaert, G. PHIRE, a deterministic approach to reveal regulatory elements in bacteriophage genomes. *Bioinformatics* **20**, 629-U699, doi:10.1093/bioinformatics/btg456 (2004).

274 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: A sequence logo generator. *Genome Research* **14**, 1188-1190, doi:10.1101/gr.849004 (2004).

275 Brewer, T. E., Stroupe, M. E. & Jones, K. M. The genome, proteome and phylogenetic analysis of *Sinorhizobium meliloti* phage Phi M12, the founder of a new group of T4-superfamily phages. *Virology* **450**, 84-97, doi:10.1016/j.virol.2013.11.027 (2014).

276 Liu, K., Linder, C. R. & Warnow, T. RAxML and FastTree: comparing two methods for large-scale maximum likelihood phylogeny estimation. *PLoS One* **6**, e27731, doi:10.1371/journal.pone.0027731 (2011).

277 Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution* **26**, 1641-1650, doi:10.1093/molbev/msp077 (2009).

278 Miller, E. S. *et al.* Bacteriophage T4 genome. *Microbiology and Molecular Biology Reviews* **67**, 86-+, doi:10.1128/mmbr.67.1.86-156.2003 (2003).

279 Koch, T., Raudonikiene, A., Wilkens, K. & Rueger, W. Overexpression, purification, and characterization of the ADP-ribosyltransferase (gpAlt) of bacteriophage T4: ADP-ribosylation of E. coli RNA polymerase modulates T4 "early" transcription. *Gene Expression* **4**, 253-264 (1995).

280 Tiemann, B. *et al.* ModA and ModB, two ADP-ribosyltransferases encoded by bacteriophage T4: Catalytic properties and mutation analysis. *Journal of Bacteriology* **186**, 7262-7272, doi:10.1128/jb.186.21.7262-7272.2004 (2004).

281 Lavigne, R., Villegas, A. & Kropinksi, A. M. in *Methods in Molecular Biology* Vol. 502 *Methods in Molecular Biology* (eds M. R. J. Clokie & A. M. Kropinski) 113-129 (2009).

282 Fuller, N. J., Wilson, W. H., Joint, I. R. & Mann, N. H. Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Applied and Environmental Microbiology* **64**, 2051-2060 (1998).

283 Zhong, Y., Chen, F., Wilhelm, S. W., Poorvin, L. & Hodson, R. E. Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. *Applied and Environmental Microbiology* **68**, 1576-1584, doi:10.1128/aem.68.4.1576-1584.2002 (2002).

284 Sullivan, M. B. *et al.* Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environmental Microbiology* **12**, 3035-3056, doi:10.1111/j.1462-2920.2010.02280.x (2010).

285 Petrov, V. M., Ratnayaka, S., Nolan, J. M., Miller, E. S. & Karam, J. D. Genomes of the T4-related bacteriophages as windows on microbial genome evolution. *Virology Journal* **7**, doi:10.1186/1743-422x-7-292 (2010).

286 Dreher, T. W. *et al.* A freshwater cyanophage whose genome indicates close relationships to photosynthetic marine cyanomyophages. *Environmental Microbiology* **13**, 1858-1874, doi:10.1111/j.1462-2920.2011.02502.x (2011).

287 Adegoke, A. A., Stenström, T. A. & Okoh, A. I. *Stenotrophomonas maltophilia* as an emerging ubiquitous pathogen: looking beyond contemporary antibiotic therapy. *Frontiers in Microbiology* **8** (2017).

288 McCutcheon, J. G., Peters, D. L. & Dennis, J. J. Identification and characterization of type iv pili as the cellular receptor of broad host range *Stenotrophomonas maltophilia* bacteriophages DLP1 and DLP2. *Viruses* **10**, doi:10.3390/v10060338 (2018).

289 Lemay, M. L., Tremblay, D. M. & Moineau, S. Genome Engineering of Virulent Lactococcal Phages Using CRISPR-Cas9. *ACS Synth Biol* **6**, 1351-1358, doi:10.1021/acssynbio.6b00388 (2017).

290 Tao, P., Wu, X., Tang, W. C., Zhu, J. & Rao, V. Engineering of Bacteriophage T4 Genome Using CRISPR-Cas9. *ACS Synth Biol* **6**, 1952-1961, doi:10.1021/acssynbio.7b00179 (2017).

291 Abedon, S. T. Lysis of lysis-inhibited bacteriophage T4-infected cells. *J Bacteriol* **174**, 8073-8080 (1992).

292 Harvey, H. *et al. Pseudomonas aeruginosa* defends against phages though type IV pilus glycosylation. *Nat Microbiol* **3**, 47-52, doi:10.1038/s41564-017-0061-y (2018).

293 Thomson, E. L. & Dennis, J. J. Common duckweed (*Lemna minor*) is a versatile high-thoughput infection model for the *Burkholderia cepacia* complex and other pathogenic bacteria. *PLoS One* **8**, e80102, doi:10.1371/journal.pone.0080102 (2013).