BACTERIOPHAGE OF *Burkholderia pseudomallei*; FRIEND OR FOE?

Thesis submitted by

Jennifer Elliman

Bachelor of Science (Hons) (JCU)
Graduate Diploma of Biomedical Sciences (JCU)
In November 2006

for the degree of Doctor of Philosophy
in the School of Veterinary and Biomedical Sciences
at James Cook University

STATEMENT OF ACCESS TO THIS THESIS

I, the undersigned, the author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere. I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and I do not wish to place any further restriction on access to this work.

J. Elliman Date

DECLARATION

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged and a list of references is given.

J. Elliman Date

ELECTRONIC COPY

I , the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library is an accurate copy of the print thesis submitted, within the limits of the technology available.

J. Elliman Date

DECLARATION OF ETHICS

The research presented and reported in this thesis was conducted within the guidelines for research ethics outlined in the Joint NHMRC/AVCC Statement and Guidelines on Research Practice (1997), The James Cook University Policy on Experimentation Ethics. Standard Practices and Guidelines (2001), and the James cook University Statement and Guidelines on Research Practice (2001). The proposed research methodology recieved clearance from the James Cook University Experimentation Ethics Review Committee (approval numbers A526 and A978).

J. Elliman Date

ACKNOWLEDGEMENTS

Thanks must firstly go to Robert Hirst and Brenda Govan for starting me off on my PhD and supervising me through a rapid learning curve. Thanks must also go to Jane Oakey for suggesting I look for bacteriophage when my initial work hit a brick wall, for being so interested in my work and for assistance with various lab protocols. Great thanks must go to my current supervisor Leigh Owens, who took me on as a student midstream following Robert's retirement, found me paid work to support myself and was always available to impart the wisdom of his years. Leigh, you saved my bacon in so many ways (and no, you're not old, even if we students have made you lose your hair). Thanks also to Jan Smith for employing me, thus keeping me financially afloat.

Thanks to the multitude of students that came and with whom I shared an office, as well as those in other offices. I'd particularly like to thank Ray for his advice in the field of molecular biology. I'd also like to thank Kjersti, for her many distractions when chapter 4 was getting overwhelming. Thank-you to Ryutaro Ueda for helping me with processing of large amounts of PCR and sequencing. Also thanks to all the academics, students, techies and office staff for their interest and help through the convoluted administrative and organisational mazes of academia.

I'd also like to thank my friends who gave me emotional support through my years of study; Kris and Simon, who fed me, listened to me babble on about things they didn't understand and made me do physical exercise; Ros who went shopping with me (lots); Brad, who proofread my thesis; and Amanda, Chris and Andrea, who not only asked me what I was doing, but took me on holidays with them.

Finally I'd like to thank my family, especially my Uncle Des & Aunty Marian, (if you don't know why, I'll tell you later). Most importantly my parents who have supported me through my whole PhD - prayerfully, emotionally, and by giving me somewhere to stay rent free! Lastly I'd like to thank Paul and Doug for encouraging me to strive by constantly asking whether I was finished yet. Thankyou all so much.

A straight line may be the shortest distance between two points, but it is by no means the most interesting - Dr Who, from 'The Time Warrior'

ABSTRACT

Lysogenic bacteriophage carrying virulence determinants have been demonstrated to be responsible for the pathogenicity of many bacteria. Bacteriophage, or components of bacteriophage, have also been successfully used in the treatment of bacterial infections. *Burkholderia pseudomallei* is the causative agent of melioidosis and has been shown to carry bacteriophage. The role of bacteriophage in virulence of *B. pseudomallei* isolates has not yet been determined, nor have bacteriophage been examined for their potential in treatment of melioidosis.

A screen for identification of bacterial isolates of interest was developed and 50 isolates were examined. Thirty-one selected isolates were then examined for bacteriophage using techniques including; transmission electron microscopy (T.E.M), mitomycin C assay, UV assay, plaque assay and restriction digestion assay. A combination of mitomycin C assay and either plaque assay or restriction digestion assay were determined to be 96.77% accurate for testing for bacteriophage in *B. pseudomallei* isolates. Five techniques for the concentration of bacteriophage (commercial Qiagen kit, magnesium hydroxide precipitation, PEG precipitation, zinc chloride precipitation, ultracentrifugation) were examined and ultracentrifugation determined to be the best. Two methods of DNA extraction (commercial nucleobond AX kit, phenol chloroform extraction) were compared and a phenol chloroform extraction was modified for use.

A bacteriophage amplification system involving inoculation of bacteriophage into a broth of host *B. pseudomallei*, followed by lysis, was developed and optimised for production of lysogenic bacteriophage of *B. pseudomallei*. Addition of a 1:1 dose of bacteriophage to bacteria at an $O.D._{600nm}$ of 0.1 in 10-100ml of broth resulted in the production of $1x10^{11}$plaque forming units (pfu)/ml of media upon lysis at 7.5 hours post-inoculation.

Lysogenic bacteriophage extracted from highly virulent *B. pseudomallei* isolate NCTC 13178 was given the name BupsΦ1 and was characterised as being from the

family *Myoviridae* with a genome 55.1kb long. This bacteriophage was then used for infection assays and molecular analysis to determine whether it played a role in virulence. Endolysin of this bacteriophage was also extracted to determine its potential for use in therapy.

Four *B. pseudomallei* isolates tested negative for the presence of bacteriophage (#13, #69, #83, E4) and one isolate of particular interest (NAFC), were infected with BupsΦ1. Bacteriophage infection was found to alter colonial morphology on Ashdown agar. Infection assays in a BALB/c mouse model were carried out and no clear relationship between addition of bacteriophage BupsΦ1 and virulence was found. One experiment with NAFC resulted in greatly increased virulence, but this could not be repeated. All other experiments where infection with bacteriophage was successful resulted in minor upregulation or downregulation of virulence. Examination of plaque production of infected and control isolates indicated that prophage stability may play a role in survival of *B. pseudomallei* as addition of bacteriophage from NCTC13178 restored lysogenic stability to NAFC in several cases.

Of the expected 55.1kb genome size from BupsΦ1, 51.3kb was sequenced with 40.9kb of this confirmed as bacteriophage. The open reading frames were determined using ORF finder and direct analysis. These open reading frames were analysed by BLASTx for putative function and several potential virulence genes were identified, as were structural, replication and lysogeny genes.

Possible virulence genes include putative anaerobic dehydrogenase and oxidoreductase genes. Putative structural genes included the terminase large subunit, portal protein, head morphogenesis, tail assembly and tail fibre genes. Putative replication and lysogeny genes included transposases, insertion elements and integrase, an RNA polymerase sigma subunit, DNA cytosine methylase, Holliday junction resolvase, repressor protein, and a weak match to *cro*, the gene responsible for triggering lysis.

Two genes of interest, the endolysin gene and a possible ADP-ribosyltransferase gene (a gene often involved in virulence) were not identified by BLASTx analysis. Techniques designed to identify genes with limited amino acid homology across species, such as identification of conserved amino acid pattern, chemo-physical comparison and phylogenetic tree analysis including bootstrap scoring, were then used to identify several open reading frames which were possible matches to these previously unidentified genes.

The endolysin of BupsΦ1 was extracted under nine combinations of conditions from literature , using a natural host system (*B. pseudomallei* #4). EDTA was found to aid lysis, while chloroform was found to have no effect. Extracts were concentrated using Centricons™ and both neat and concentrated extracts were tested for their ability to lyse both killed and live *B. pseudomallei* #4 in broth and plate format.

Neither the extracted endolysin nor its concentrate was found to lyse any of the *B. pseudomallei* in a form not attributable to live bacteriophage. Hence endolysin was determined not to function "from without" against *B. pseudomallei*. As such, this possibility for treatment of *B. pseudomallei* was eliminated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aa	amino acid
ADP	adenosine diphosphate
ADP-RT	ADP- ribosyltransferase
BHIB	brain heart infusion broth
bp	base pair
BupsΦ1	*Burkholderia pseudomallei* bacteriophage 1, from isolate NCTC 13178
BV	bacterial vaginosis
C14-NAD	carbon-14 nicotinamide adenine dinucleotide
cos	cohesive
cfu	colony forming unit
CIAP	calf intestinal alkaline phosphatase
CTAB	hexadecyltrimethyl ammonium bromide
ddH_2O	deionised water
DNA	deoxyribonucleic acid
DNase 1	deoxyribonuclease 1
dNTP	PCR nucleotide mix (deoxynucleotide triphosphate)
DTT	dithiothreitol
ds	double stranded
EDTA	ethylenediaminetetraacetic acid
EF-2	elongation factor two
EHEC	enterohaemorrhagic *E. coli*
EPEC	enteropathogenic *E. coli*
g	gravity
GI	genetic island
ICTVdB	The Universal Database of the International Committee on Taxonomy of Viruses
ID_{50}	50% endpoints of infectious dose
int	integrase
IPTG	isopropyl-β-D-thiogalactopyranoside
JCU	James Cook University
kb	kilobase
LB	Luria Bertani
Mb	megabase
mwt	molecular weight
N.P.V.	negative predictive value
O.D.	optical density

ORF	open reading frame
pac	packaging
PCR	polymerase chain reaction
PBS	phosphate buffered saline
PC3	physical containment level three
PEG	polyethylene glycol
PFGE	pulse field gel electrophoresis
pfu	plaque forming unit
PI	pathogenicity island
P.P.V.	positive predictive value
rpm	revolutions per minute
RNA	ribonucleic acid
RNase A	ribonuclease A
SBA	sheep blood agar
SDS	sodium dodecyl sulphate
SLT	shiga-like toxin
STS	serine-threonine-serine
TAE	tris acetate
TBE	tris borate
T.E.M.	transmission electron microscopy
Tm	melting temperature
tRNA	transfer ribonucleic acid
TSB	tryptone soya broth
TTSS	type three secretion system
UV	ultraviolet
UV-C	ultraviolet radiation at 254nm
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
Xis	excisase

CHAPTER 1
GENERAL INTRODUCTION

Bacteriophage are viruses that infect bacteria (Anon, 1921). Of those examined by transmission electron microscopy (T.E.M.), approximately 96% belong to the order Caudovirales. These are double stranded DNA viruses consisting of the families *Myoviridae, Syphoviridae* and *Podoviridae* (Maniloff and Ackermann, 1998). Bacteriophage play a role in the ecosystems in which they exist, by both controlling biomass and make-up of bacterial populations by lysis (Fuhrman, 1999; Kilic *et al.*, 2001) and by altering bacterial genomes and hence characteristics by lysogeny (Boyd, 2004).

Lysis of bacteria by bacteriophage is of interest in the treatment of disease. Bacteriophage therapy, whereby cocktails of bacteriophage are used either in isolation or in combination with antibiotics, has been successfully used in both human and animal treatment (Smith *et al.*, 1987; Gorski and Weber-Dabrowska, 2005; Jikia *et al.*, 2005). Bacteriophage proteins, including endolysin, have also been examined for their use as a treatment for bacterial diseases of humans and plants (Nelson *et al.*, 2001; Kim *et al.*, 2004). The use of bacteriophage proteins rather than whole bacteriophage eliminates the possibility of bacteriophage entering a lysogenic lifecycle in the bacteria and potentially increasing virulence.

Virulence factors carried on bacteriophage include genes expressing extracellular toxins, antigenic alteration proteins, effector proteins involved in invasion, enzymes required for intracellular survival, serum resistance from outer membrane proteins and adhesins for bacterial host attachment as well as a range of other functions (Bensing *et al.*, 2001; Boyd and Brussow, 2002).

Burkholderia pseudomallei is a Gram-negative, facultative intracellular bacterium (Jones *et al.*, 1996) which is the causative agent of melioidosis (Dance, 2000). Melioidosis is a potentially lethal disease, with symptoms of infection ranging from an asymptomatic carrier state to septicaemia resulting in death within 48 hours (White, 2003). *B. pseudomallei* is also considered a potential bioterrorism tool

(Bossi *et al.*, 2004; Cheng *et al.*, 2005a).

Some of the identified virulence factors of *B. pseudomallei* have homologues in other bacteria which have been identified as being bacteriophage encoded. These include an ADP-ribosylating exotoxin (Mohamed *et al.*, 1989a; Mohamed *et al.*, 1989b) and putative adhesins (Holden *et al.*, 2004). The possibility of bacteriophage involvement in the virulence must be considered.

Several studies (Leclerc and Sureau, 1956; Denisov and Kapliev, 1991; Denisov and Kapliev, 1995) have found bacteriophage infecting *B. pseudomallei*. Examination of *B. pseudomallei* genomes by subtractive hybridisation (DeShazer, 2004) has also identified the presence of a prophage in a *B. pseudomallei* isolate. Where examined by T.E.M., these bacteriophage were found to belong to the families *Myoviridae* and *Syphoviridae* (Denisov and Kapliev, 1995; DeShazer, 2004).

Only two complete *B. pseudomallei* bacteriophage genomes have been reported in GenBank (http://www.ncbi.nlm.nih.gov/) and no experiments regarding either use of *B. pseudomallei* bacteriophage in treatment of infection, or virulence assays have been reported.

The aims of this research are to examine a library of *Burkholderia pseudomallei* isolates at James Cook University for the presence of bacteriophage and to further examine isolates of interest. Bacterial isolates of low virulence showing no presence of bacteriophage will be infected with lysogenic bacteriophage from an isolate of high virulence. These infected isolates will be examined to determine whether any changes in virulence occur. In addition, sequencing of the bacteriophage genome used for virulence assay will be carried out and the genome analysed. This will broaden our understanding of virulence and add to the limited amount of available sequence data on *B. pseudomallei* bacteriophage. The possibility of using the endolysin of this bacteriophage to kill *B. pseudomallei* will also be examined. Although lysogenic bacteriophage are not considered to be good candidates for phage therapy, their component proteins may still be of use.

CHAPTER 2
LITERATURE REVIEW

2.1 The History of Bacteriophage

In 1896 British chemist Ernest Hanbury Hankin discovered that the waters of the Ganges River could destroy cholera bacteria (Parfitt, 2005), but it was not until 1915 that Frederick Twort first observed what he believed might be viruses that could infect bacteria (Twort, 1915). In 1917 Felix d'Herelle independently determined this and named these viruses bacteriophage, meaning "eaters of bacteria" (Anon, 1921). Twort and d'Herelle are generally considered the discoverers of bacteriophage, though some researchers have queried the independent nature of d'Herelle's discovery (Duckworth, 1976).

Research in subsequent decades continued to examine the nature of these viruses and they were used as model systems to investigate many wider aspects of virology, such as viral structure, replication and genetics. The experiments that determined that DNA rather than protein was the carrier of genetic information were carried out using T2 and T4 bacteriophage (Hershey and Chase, 1952). The first gene mapping was carried out on the genome of T4 (Benzer, 1955), which was also used as a tool to demonstrate discontinuous replication of DNA (Okazaki *et al.*, 1968). Restriction endonuclease modification systems were first detected using bacteriophage including λ (Bertani and Weigle, 1953; Nathans and Smith, 1975). Bacteriophage λ has been extensively used for a range of objectives, including development of an understanding of gene regulation (Ptashne *et al.*, 1982) and for use as a vector to analyse genes from a wide range of other organisms (Morrow, 1979; Sambrook and Russell, 2001). In addition, the rapid accumulation of information on bacteriophage genome sequence has provided insights into the origins of infectious disease as well as identifying novel biochemical mechanisms (Karam, 2005).

Intensive studies on the therapeutic use of bacteriophage for treatment of infectious diseases began in 1920 (Mathur *et al.*, 2003). Following the discovery of penicillin in 1928 (Bentley, 2005), study into the therapeutic possibilities of bacteriophage was in essence abandoned in favour of the wider ranging usage of antibiotics (Merril *et*

al., 1996; Mathur *et al.*, 2003). Research on bacteriophage therapy remained almost solely the responsibility of Soviet science (Chanishvili *et al.*, 2001; Verthe *et al.*, 2004). The lack of international peer review and a limited number of English language articles has meant that much of this work was unavailable or ignored in the west (Sulakvelidze, 2005).

Interest in bacteriophage therapy was rekindled in the west in recent years primarily due to the advent of antibiotic resistance in bacteria (Mathur *et al.*, 2003; Huff *et al.*, 2004). In former Soviet Georgia, bacteriophage therapy is well advanced, with treatment routinely used (Parfitt, 2005) and much Soviet literature reporting the use of bacteriophage to treat infection or as a prophylactic existing from the 1930's onward (Chanishvili *et al.*, 2001).

2.2 Definitions

There are various terms which are used to describe bacteriophage and bacteriophage behaviours, some of which are used interchangeably. **Lysogens** can be defined as bacteria that harbour prophage. Bacteriophage are called **prophage** when the viral genome is integrated in the bacterial genome. Bacteriophage that can do this are defined as **temperate** or **lysogenic**. **Superinfection immunity** is the resistance to further infection by a similar or the same bacteriophage conferred by bacteriophage already infecting the bacteria (Kilic *et al.*, 2001). **Lysogeny** can be described as the perpetuation of prophages as part of the bacterial replicating system (Barksdale and Arden, 1974). **Pseudolysogeny** has been described as a carrier state in which bacteriophage DNA is not integrated into the host genome, but does not cause immediate lysis, continuing within the host system through multiple generations, but not in all colonies (Baess, 1971). It can be caused by starvation of the host cell, limiting the ability of the bacteriophage to replicate (Miller and Ripp, 1998). Lysogens can be induced to produce lytic viruses (Fuhrman, 1999). **Lysis** is the destruction of the host bacterium to release progeny bacteriophage into the environment (Young, 1992). Bacteriophage that do not enter a lysogenic state are **obligatorily lytic** or **virulent** (Madsen *et al.*, 2001).

2.3 Geography

Bacteriophage are found in almost all environments on Earth, from the depths of the ocean to hot springs, and can be isolated from almost any material that will support bacteria (Dabrowska *et al.*, 2005). There is evidence that the diversity of bacteriophage is about an order of magnitude higher than that of bacteria (Weinbauer and Rassoulzadegan, 2004) which has implications in the classification of bacteriophage.

2.3.1 Environmental significance

Bacteriophage can have a great effect on both the biomass and the characteristics of bacteria. In marine environments, viruses can limit bacterial abundance to several orders of magnitude below resource limits (Fuhrman, 1999). The lysis of bacteria by bacteriophage results in an increase in inorganic nutrients (Fuhrman, 1999) as well as increasing dissolved organic matter (Paul *et al.*, 2002) and can alter bacterial ecologies (Wommack and Colwell, 2000). In one study, the biomass of a *Pseudomonas fluorescens* biofilm was reduced by 85% after infection with the lytic bacteriophage ΦS1 (Sillankorva *et al.*, 2004). Examination of the bacterial and bacteriophage ecology of the human vagina has shown evidence that *Lactobacillus* bacteriophage could be a factor in development of bacterial vaginosis (BV), by lysis of favourable *Lactobacillus* allowing other less favourable bacteria to multiply. It was theorised that transmission of *Lactobacillus* lysogens was the reason that BV is associated with sexual transmission (Kilic *et al.*, 2001).

Lysogeny plays a role in bacterial ecologies with high percentages of bacteria determined to be lysogens. Over 40% of marine bacterial isolates carried inducible temperate bacteriophage (Jiang and Paul, 1998). Bacteriophage are not, as noted previously, limited to aquatic environments. Other screening experiments have shown 27 of 34 isolates of *Salmonella enteritidis* (Jacob and Wollman, 1959), 83 of 107 *Escherichia coli* isolates, 136 of 173 *Salmonella enterica* isolates and 38 of 68 dairy *Streptococcus* strains (Casjens, 2003) to be lysogens.

The viral genome may carry genetic elements that can, once inserted in the bacterial genome, provide an advantage to the bacteria. This horizontal transfer of genetic information by bacteriophage, also known as transduction, has been recognised as playing a major role in microbial diversity, the development of niche organisms, and changes in pathogen-host interactions (Fuhrman, 1999; Miao and Miller, 1999; Fuhrman and Schwalbach, 2003; Weinbauer and Rassoulzadegan, 2004). It has been estimated that marine bacteriophage transduce 10^{28} base pairs of DNA per year in the world's oceans (Paul *et al.*, 2002). In many cases, bacteriophage have been found to encode genes which give the bacteria a survival advantage by altering structural characteristics (Pruzzo *et al.*, 1980; Mirold *et al.*, 2001), resistance to antibiotics (Paul *et al.*, 2002) or increased virulence (Boyd and Brussow, 2002). This will be further covered in detail in Section 2.7; Virulence Factors.

2.4 Taxonomy

The initial classification system for bacteriophage was designed in the 1930's and was based on different host specificities of various bacteriophage (Nelson, 2004). With the advent and use of the electron microscope in the 1940's and 50's, bacteriophage were classified based on morphotype (Luria *et al.*, 1943). They have since been further classified with respect to molecular characteristics as techniques to identify DNA or RNA type and replication strategies have developed (Ackermann *et al.*, 1978; Maniloff and Ackermann, 1998). The current typing system identifies at least 13 distinct groups of bacteriophage (Maniloff and Ackermann, 1998), including those infecting Archaea. These are very diverse structurally and genetically and those bacteriophage infecting Archaea will not be discussed further. Of those bacteriophage examined by electron microscopy, about 96% belong to the order Caudovirales (Table 2.1).

Table 2.1 Taxonomy of bacteriophage. Viruses infecting only Archaea are not included in this list - taken from the following sources (Ackermann, 1998; Maniloff and Ackermann, 1998; Dabrowska *et al.*, 2005) http://www.ncbi.nlm.nih.gov/ICTVdb/ (ICTVdB The Universal Database of the International Committee on Taxonomy of Viruses)

Order (where assigned)	Family or Group	Genus	Type member	Particle Morphology	Genome
Caudovirales	*Myoviridae*	*T4-like viruses*	*Enterobacteria phage T4*	contractile tailed phage	linear ds DNA
		P1-like viruses	*Enterobacteria phage P1*		
		P2-like viruses	*Enterobacteria phage P2*		
		Mu-like viruses	*Enterobacteria phage Mu*		
		SP01	*Bacillus phage SP01*		
	Podoviridae	*T7-like viruses*	*Enterobacteria phage T7*	short tailed phage	linear ds DNA
		Φ29-like viruses	*Bacillus phage Φ29*		
		P22-like viruses	*Enterobacteria phage P22*		
	Siphoviridae	*λ-like viruses*	*Enterobacteria phage λ*	non-contractile tailed phage	linear ds DNA
		T1-like viruses	*Enterobacteria phage T1*		
		T5-like viruses	*Enterobacteria phage T5*		
		c2-like viruses	*Lactococcus phage c2*		
		L5	*Mycobacterium phage L5*		
	Cystoviridae	*Cystovirus*	*Ø6*	isometric	3 segments ds RNA
	Inoviridae	*Inovirus*	*Enterobacteria phage M13*	rod	circular ss DNA
		Plectrovirus	*Acholeplasma phage MV-L51*		
	Leviviridae	*Levivirus*	*Enterobacteria phage MS2*	icosahedral	1 (+)strand RNA
		Allolevirus	*Enterobacteria phage Qβ*		
	Microviridae	*Microvirus*	*Enterobacteria phage ØX174*	icosahedral	circular ss DNA
		Spirovirus	*Spiroplasma phage 4*		
		Bdellomicrovirus	*Bdellomicrovirus phage MAC1*		
		Chlamydiamicrovirus	*Chlamydia phage 1*		
	Plasmaviridae	*Plasmavirus*	*Acholeplasma phage L2*	pleiomorphic	circular ds DNA
	Tectiviridae	*Tectivirus*	*Enterobacteria phage PRD1*	icosahedral	linear ds dna
	Corticoviridae	*Corticovirus*	*Alteromonas phage PM2*	icosahedral	circular ds DNA

The nomenclature of bacteriophage remains confused, with the use of numerals, the Greek letter Φ and the Latin letter P frequently used (Ackermann *et al.*, 1978). There are also cases of unrelated bacteriophage being assigned similar names due to dissimilar naming protocols. For example, a bacteriophage of *Pseudomonas aeruginosa* and a bacteriophage of *Vibrio cholerae* have the names ΦCTX (Hayashi *et al.*, 1990) and CTXΦ (Waldor and Mekalanos, 1996) respectively. Ackermann (1978) suggested the use of the first two letters of the genus of the bacterial host, followed by the first, or first two letters, of the species name, followed by numbers, Latin capitals and/or Greek letters when naming bacteriophage.

Our current ability to sequence entire genomes and proteomes may, in the future, result in a new classification system as some bacteriophage which are morphologically dissimilar are genomically very similar and *vice versa* (Rohwer and Edwards, 2002). In addition, bacteriophage genomes (prophage) are being sequenced within bacterial genomes, without ever identifying a mature virion (Nelson, 2004). However, since it is estimated that we have sampled less than 0.0002% of the global bacteriophage genomes (Rohwer, 2003), a genomic classification system may be some distance away.

2.5 Structure of Caudovirales Bacteriophage

As 96% of bacteriophage identified thus far are of the Order Caudovirales (Maniloff and Ackermann, 1998), further discussion will be restricted to this Order. The Caudovirales are tailed viruses,with the name Caudovirales being derived from the Latin *cauda*, meaning tail. These bacteriophage have a head, or capsid, in which DNA is stored, and a tail. Variation within this basic structure is one of the determinants of taxonomic separation (Ackermann, 1998).

2.5.1 The capsid

The capsid is icosohedral in shape, with rare elongated variations. It appears smooth under electron microscopy and ranges in diameter from 34 to 160nm with a majority at 60nm. Capsomers (the morphological subunits) are also present (Bradley, 1967; Ackermann, 2003). The family *Myoviridae* are generally larger than both other

families (Ackermann, 1998).

2.5.2 The tail

The tail is structured of a hollow tube of fixed length and width, built of stacked rows of subunits and generally has a six-fold symmetry (Ackermann, 1998). Members of the family *Siphoviridae* have long non-contractile, flexible tails, the family *Podoviridae* have short variants on this, while the family *Myoviridae* have long, rigid, contractile tails (Ackermann *et al.*, 1992). The family *Myoviridae* tails consist of a tail tube surrounded by a sheath, separated from the head by a neck. They are of sixfold symmetry, with subunits arranged in helix format. On contraction, these subunits slide over each other, forming a short cylinder (Ackermann, 1998). Tail lengths can vary widely, but are typically conserved within a species. A ruler protein has been identified in some bacteriophage and this acts as a tape-measure around which tail tube monomers polymerize. Alterations to the ruler protein will alter the length of the tail tube (Katsura and Hendrix, 1984).

2.5.3 Other structures

There is a small disk located inside the head at the site of tail attachment which is known as the connector. The connector holds the head and tail together and has functions in head assembly and DNA encapsidation (Ackermann, 1998). Tailed bacteriophage can also have base plates, tail spikes and tail fibres, though the number and shape of these can vary (Ackermann, 1998). These have various functions in infection of bacteria and will be described later (2.6.1 Attachment; 2.6.2 Penetration).

2.5.4 Genomic structure

Genomes of Caudovirales consist of linear double stranded (ds) DNA (http://www.ncbi.nlm.nih.gov/ICTVdb/ ; ICTVdB The Universal Database of the International Committee on Taxonomy of Viruses). They range from 17kb to in excess of 700kb in length (Ackermann, 2003) with a majority of genomes being about 50kb (Ackermann, 1998). Some genomes contain cohesive (*cos*) sites near either end allowing circularisation of the genome after infection. Packaging of DNA can be either of a single genome, or by a headful mechanism, where the genome is

continually copied into the capsid until it is full (Streisinger *et al.*, 1967; Ackermann, 1998). DNA may be concatomeric (head-to-tail repeats of a sequence) or unit length prior to packaging (Ackermann, 1998) with concatomeric DNA formed from recombination between linear DNAs or rolling circle replication (Maniloff and Ackermann, 1998). Cleavage of concatomeric DNA for packaging can occur at; a) unique sites resulting in blunt termini or *cos* ends with packaging starting and finishing at a *cos* site, b) *pac* sites (sequences recognised by terminase complex) to produce DNA molecules with limited circular permutation and terminal redundancy (excess coding DNA at the terminal end) and where packaging starts at the *pac* site and continues until the head is full or c) random sites to produce circularly permuted, terminally redundant DNA (Ackermann, 1998; Maniloff and Ackermann, 1998).

From comparisons of bacteriophage genomes, it is apparent that bacteriophage genomes have a mosaic nature, with gene order not conserved between species (Chopin *et al.*, 2001; Hertveldt *et al.*, 2005). The following observations of gene order can be made. Genes with related function generally cluster together, though non-structural gene order does not follow any general pattern. Structural genes are generally separate from other genes and, of these, head genes precede tail genes (Casjens, 2003).

2.6 Lifecycle

Bacteriophage use several strategies for replication within bacteria and can exist outside of bacteria for some time, dependant on temperature and levels of organic matter present (Wommack and Colwell, 2000). Bacteriophage attach to the cell wall of bacteria, insert their DNA and then primarily enter either a lytic or lysogenic lifecycle (Figure 2.1). Some bacteriophage have been identified as being obligatorally lytic (or virulent) and will kill the bacterium they have invaded in order to replicate (Madsen *et al.*, 2001). Others can enter a lysogenic cycle where the bacteriophage DNA will be replicated along with bacterial DNA during bacterial replication. These bacteriophage can enter the lytic cycle when an appropriate trigger, such as environmental stressor, is identified (Weinbauer and Suttle, 1996). Some bacteriophage have also been identified as being lysogenic in some strains of

an organism, but lytic in others (Cluzel *et al.*, 1987; Kilic *et al.*, 2001). Furthermore, some bacteriophage can enter either a lytic or lysogenic lifecycle depending on environmental conditions such as temperature or multiplicity of infection (Bertani and Nice, 1954). Bacteria have also been reported to be polylysogenic, carrying multiple bacteriophage in their genomes (Groman and Laird, 1977; Denisov and Kapliev, 1991).

Figure 2.1 Lytic and lysogenic lifecycle of bacteriophage using bacteriophage λ as the example. Taken from http://www.accessexcellence.org/RC/VL/GG/bact_Lambda.html, (Alberts *et al.*, 1998)

2.6.1 Attachment

Bacteriophage of the Order Caudovirales adsorb to the bacterial cell wall, flagella or pili (Bradley, 1967) via tail fibres. These tail fibres absorb to receptors on the cell surface (Weitz *et al.*, 2005). Examples of receptors include LamB, used by λ (Thirion and Hofnung, 1972) and lipopolysaccharide receptors, used by T4 (Leiman *et al.*, 2004).

Alteration in surface structures can alter sensitivity to bacteriophage. Some models of co-evolutionary dynamics (Weitz *et al.*, 2005) suggest that as there are conformational changes in cell surface receptors on bacteria, there will also be conformational changes in tail fibres to deal with this. Other models describe an imperfect lock-key mechanism where the match is not required to be perfect. Some alteration to LamB, the receptor protein for λ and some other bacteriophage (Charbit and Hofnung, 1985), can be undertaken without compromising successful attachment of λ (Gehring *et al.*, 1987).

The host range for bacteriophage can vary with most bacteriophage infecting certain species or even strains of bacteria (Rohwer, 2003). In one study, a range of toxin encoding bacteriophage extracted from *E. coli* O157:H7 isolates were found to have variable *E. coli* host ranges including non O157 serotype *E. coli* hosts (Gamage *et al.*, 2004). In another study, the HP1/S2 group of *Haemophilus influenzae* bacteriophage lysogenised only Rd strains of *H. influenzae* (Williams *et al.*, 2002). Host ranges spanning species are rare, although examples have been reported. Many bacteriophage from lactobacilli have a broad host range within the genus *Lactobacillus* including the capacity to superinfect these hosts, but some bacterial strains remained uninfected, implying variation in phage receptor might be responsible (Kilic *et al.*, 2001; Lu *et al.*, 2003). Similarly *Burkholderia mallei* is sensitive to bacteriophage from *Burkholderia pseudomallei* (Smith, 1957).

2.6.2 Penetration

After absorption of tail fibres to the cell wall, the baseplate comes into contact with the cell surface receptors and, in the case of T4 of the family *Myoviridae*,

experiences a conformational change from hexagon to star configuration that causes the tail sheath to contract (Crowther *et al.*, 1977; Crawford and Goldberg, 1980). This drives the tail tube through the bacterial cell wall and the periplasmic peptidoglycan layer is digested in a localised manner by a lysozyme domain at its tip (Leiman *et al.*, 2004). Subsequently the bacteriophage DNA is released into the cell through the tube (Leiman *et al.*, 2004). The families *Syphoviridae* and *Podoviridae* do not have contractile tails (Ackermann, 2003) but do have baseplates and tail fibres that act in a similar fashion, with at least some bacteriophage having tail fibres responsible for peptidoglycan digestion for DNA insertion (Vegge *et al.*, 2005). Bacteriophage proteins involved with injection have been identified (Bryant and King, 1984) and structural changes of proteins binding to DNA are believed to be involved in triggering injection (Ilag *et al.*, 1993).

Once bacteriophage have adsorbed to the cell wall, bacteria have several defences against bacteriophage infection. These include blocking DNA from being injected (Scandella and Arber, 1974), digestion of bacteriophage DNA (Bourniquel and Bickle, 2002) and cell death to prevent bacteriophage spread (Hazan and Engelberg-Kulka, 2004). Bacteriophage can evade DNA targeted defences by techniques including blocking production of restriction enzymes and methylation or unusual base substitution of bacteriophage DNA, so restriction enzymes do not recognise them (Kruger and Bickle, 1983).

2.6.3 Replication: lysis or lysogeny

Once DNA has penetrated the cell wall, bacteriophage will either enter a lytic replication cycle, in which the bacteriophage is replicated without insertion into the host chromosome, or a lysogenic cycle. In a lysogenic cycle the DNA either maintains itself as a plasmid and replicates autonomously (Rosner, 1972) or integrates into the chromosome of the bacteria in which case the bacteriophage is now called a prophage. A minimum number of genes are expressed to maintain this state. Genes which are expressed at this time include those which suppress replication genes and those which confer superinfection immunity (Clark *et al.*, 1991; Clapper *et al.*, 2004; Chen *et al.*, 2005).

2.6.3.1 Lysis

The lytic cycle can be simply described as follows. Transcription of RNA, usually by the hosts' RNA polymerase, begins almost immediately after DNA enters the cell. Only those bacteriophage genes that have promotors recognised by the host are transcribed. These are called early genes and are mostly involved with DNA synthesis. They include genes encoding DNA polymerase, primase, ligase and helicase, which aid in the replication of bacteriophage DNA (Snyder and Champness, 1997). The late genes, which have promoters not recognised by the host RNA polymerase alone are transcribed next. These include genes encoding for assembly of the head and tail (Snyder and Champness, 1997) and genes for lysis of the cell. Although there are exceptions (Young, 1992; Bernhardt *et al.*, 2002; Loessner, 2005), in general there are two phage induced factors that contribute to the lysis of the cell wall. Holin forms a hole in the cell membrane to allow the enzyme endolysin access to the peptidoglycan and the endolysin degrades the cell wall peptidoglycan (Young, 1992; Wang *et al.*, 2000). This causes osmotic lysis of the bacterial cell and release of the progeny bacteriophage (Loessner, 2005). From ten to 500 phages may be produced per infected bacterium (Wommack and Colwell, 2000). The entire process, from penetration of DNA to lysis of the cell takes less than one hour for many bacteriophage (Ackermann, 1998).

In reality, the replication of bacteriophage can be more complex than this with intermediate stages. Genes can be described as immediate early, early, delayed early, middle and late, with multiple layers of regulation and timing of replication (Snyder and Champness, 1997). Expression of different genes can be regulated by various specific mechanisms. One or more gene products synthesised at one stage of development can turn on synthesis of the next stage, they can also turn off transcription of the previous stages genes. Of note, bacteriophage T7 carries a phage encoded RNA polymerase (Snyder and Champness, 1997). Once decoded, this RNA polymerase transcribes the late genes. The promoters for the late genes are only recognised by the phage RNA polymerase, ensuring they are not decoded earlier (Snyder and Champness, 1997). Phage T4 late genes have also been determined to have promoters which are not recognised by the host RNA polymerase. In this case

however, it was determined that a phage encoded sigma factor was produced (Miller *et al.*, 2003) which binds to the host RNA polymerase, changing its specificity so that it will only recognise the bacteriophage promoters. This second method of transcriptional control is not uncommon, with many other bacteriophage determined to transcribe alternate sigma factors. Phage T4 also has a second level of transcriptional control of the late genes. The RNA polymerase will not function unless the phage DNA is actively replicating, with a protein binding to both the replication fork and the RNA polymerase as the trigger to activate the RNA polymerase (Snyder and Champness, 1997; Miller *et al.*, 2003). This ensures that structural genes are not expressed without copies of DNA already present. Neither T4 nor T7 cyclize, with T7 replicating by producing concatomers by pairing between complementary ends of DNA and T4 producing concatomers by recombination of repeated sequences at the DNA ends (Snyder and Champness, 1997).

2.6.3.2 Lysogeny

In lysogeny, the bacteriophage genome is integrated into the bacterial genome and remains integrated until lysis is triggered. Therefore, the bacteriophage genome can be replicated with each replication of the bacterial genome (Kokjohn, 1989). Integration usually involves site-specific recombination between short sequences on the bacterial chromosome and the bacteriophage genome (Cheetham and Katz, 1995). Bacteriophage may integrate within a single bacterial attachment site, as is the case with ΦCTX of *Pseudomonas aeruginosa* which binds to the $tRNA^{ser}$ gene (Hayashi *et al.*, 1993) or multiple sites, as is the case with bacteriophage MAV1 of *Mycoplasma arthritidis* (Voelker and Dybvig, 1998). The use of tRNA genes as insertion sites is common to a diverse range of bacteriophage as tRNA genes are present in multiple copies in the bacterial genome, providing multiple possible integration sites (Cheetham and Katz, 1995). A few temperate phage, such as Mu have a transposition site and can move from site to site in the bacterial chromosome (Ackermann, 1998). They also use replicative transposition to produce multiple copies of themselves (Summer *et al.*, 2004).

Environmental factors can trigger conversion from a lysogenic state to a lytic one, most often by triggering bacterial DNA repair mechanisms, which include the protein recA. This cleaves the repressor maintaining the lysogenic state and lysis begins (Weinbauer and Suttle, 1996). Reported triggers include ultraviolet radiation (Kidambi *et al.*, 1996), mitomycin C (Weinbauer and Suttle, 1996), benzo[α]pyrene diol epoxide (a chemical in cigarette smoke) (Kilic *et al.*, 2001), quinolones (Zhang *et al.*, 2000) and neutrophil activating products such as hydrogen peroxide (Wagner *et al.*, 2001).

2.6.4 Molecular replication of bacteriophage λ

The replication of bacteriophage λ has probably been studied at the molecular level more than any other bacteriophage. As it may be lytic or enter a lysogenic state (Figure 2.1), it is useful as an example to describe the replication of bacteriophage. It is a linear, double stranded DNA bacteriophage with 12 base single stranded *cos* ends. These can be paired and joined with DNA ligase to form closed circular DNA molecules, in which format the DNA replicates (Sambrook and Russell, 2001).

After the DNA has circularised, immediate early transcription is initiated at two promotors, p_L and p_R, on either side of the *cI* (repressor) gene (Ptashne and Gilbert, 1970). This transcription is mediated by *E. coli* RNA polymerase and terminates at the ends of the *N* and *cro* genes respectively, though transcription can carry on through the *cro* gene to t_{R2} (Sambrook and Russell, 2001). The product of *N* (pN) encodes an antiterminator protein which permits host RNA polymerase to read through the transcriptional terminators t_L and t_{R1} into the delayed early genes *cII* and *cIII* which are necessary for the lytic growth of this bacteriophage (Snyder and Champness, 1997). This protein is unstable and requires continuous production to express the proteins CII and CIII. Transcription of *N* is negatively controlled by the proteins CI and Cro and pN may negatively autoregulate its own translation (Sambrook and Russell, 2001).

The CII protein is the activator of λ genes that repress lytic functions and catalyse integration of the viral DNA into the host chromosome (Sambrook and Russell,

2001). The decision between lysogeny and lysis is influenced by the multiplicity of infection and the nutritional state of the cell. The higher the multiplicity of infection and the worse the nutritional state, the higher the frequency of lysogeny (Sambrook and Russell, 2001).

The CII protein is constantly degraded by proteases present in the cell. If levels of CII protein remain below a threshold level, phage undergoes lytic replication. This is the sequence of events which occurs in the vast majority of cells which become infected (Sambrook and Russell, 2001). The natural rate of degradation of the CII protein can be enhanced or diminished by several bacterial gene products including the CIII protein (Oppenheim *et al.*, 2005). Where the concentration of CII protein builds up, transcription of *cI* is enhanced and intracellular levels of the CI repressor protein rise. The repressor binds to the o_R and o_L operators, which prevents transcription of most phage genes including *cro* (Oppenheim *et al.*, 2005). The level of CI protein is maintained by both positive and negative feedback mechanisms, since at high concentrations the repressor also binds to the left-hand end of o_R and prevents transcription of *cI* (Oppenheim *et al.*, 2005).

Two proteins, the bacteriophage encoded integrase (int), production of which is triggered by CII protein (Oppenheim *et al.*, 2005) and the host encoded integration host factor are required for integration of bacteriophage DNA into the host chromosome. Int is a type one topoisomerase which cuts and rejoins DNA strands one at a time. Int generates a Holliday structure at *attP*, a ~240bp region of bacteriophage DNA with a 15bp core which is identical between bacteriophage and host chromosomes, and attB a 21bp region on the host chromosome. Once integrated, the inserted prophage is bracketed by 15bp repeats in different orientation. During excision, using the Int and Xis (excisase) proteins, these sites regenerate the *attP* and *attB* sites and single stranded circular bacteriophage DNA is produced (Sambrook and Russell, 2001).

The autoregulation of CI protein synthesis keeps the cell in a stable state of lysogeny (Ptashne *et al.*, 1982). To trigger lysis from this state, a host cell protein RecA is

necessary. Physiological stresses such as exposure of the cell to ultraviolet irradiation or DNA damaging materials results in the induction of RecA (Kokjohn, 1989). This protein cleaves the CI repressor protein (Weinbauer and Suttle, 1996).

To prevent the cell re-entering the lysogenic state, *cro* is transcribed from PR when the repressor protein is not bound to o_R. The Cro protein also binds to o_R, but unlike the CI protein, which preferentially binds to the right-hand end of o_R, the Cro protein binds preferentially to the left-hand end of o_R, preventing the transcription of *cI* and enhancing its own transcription in a positive feedback loop (Ptashne *et al.*, 1982). This locks the phage into a lytic cycle and production of progeny bacteriophage and lysis of the bacterial cell occurs.

2.7 Virulence Factors

Transduction is the transfer of bacterial genes among bacteria mediated by bacteriophage (Kokjohn, 1989) and is an important mechanism by which bacteria can acquire virulence factors (Miller, 2001). Genes which play a role in bacteriophage replication (Bensing *et al.*, 2001) or superinfection immunity (Clark *et al.*, 1991; Normark *et al.*, 2005) can also become bacterial virulence factors and the insertion position of the bacteriophage can cause phenotypic changes and silence or trigger gene expression (Coleman *et al.*, 1989).

Virulence factors which are encoded on bacteriophage can form various functions. Those identified to date include genes expressing extracellular toxins, antigenic alteration proteins, effector proteins involved in invasion, enzymes required for intracellular survival, serum resistance from outer membrane proteins and adhesins for bacterial host attachment as well as a range of other functions (Bensing *et al.*, 2001; Boyd and Brussow, 2002).

Virulence is associated with the presence of bacteriophage in *V. cholerae* (Waldor and Mekalanos, 1996; Faruque *et al.*, 1998). The toxin causing the rice-water stool of diarrhoeal cholera is encoded from a lysogenic bacteriophage CTXΦ, integrated into the larger of the two *V. cholerae* chromosomes (Blakely, 2004).

Other examples of bacteria which have acquired virulence factors from temperate bacteriophage include *Corynebacterium diphtheriae*, in which the diphtheria toxin gene, *tox*, is encoded on beta corynephage (Freeman, 1951; Freeman and Morse, 1952; Laird and Groman, 1976b) and *Streptococcus pyogenes*, in which the type A streptococcal exotoxin gene, *speA*, is encoded on bacteriophage T12 (Johnson *et al.*, 1986; Weeks and Ferretti, 1986). The *ctx* gene encoding the cytotoxin of *P. aeruginosa* is carried on ΦCTX (Hayashi *et al.*, 1993). The shiga toxin produced by *Shigella sonnei* is encoded on a bacteriophage, (Strauch *et al.*, 2001). *E. coli* can also express shiga-like toxins (SLT) (Betley *et al.*, 1986). The genes for these are found on a variety of bacteriophage and these toxins are biologically similar but antigenically different (Scotland *et al.*, 1983; Newland *et al.*, 1985; Strockbine *et al.*, 1986; Gamage *et al.*, 2004). In addition, bacteriophage ΦC3208 and other temperate bacteriophage have been found to cause expression of enterohemolysin in *E. coli* (Beutin *et al.*, 1993). A single bacteriophage can also cause the expression or repression of several virulence factors, with serotype F bacteriophage of *Staphylococcus aureus* capable of mediating the simultaneous triple-lysogenic conversion of enterotoxin A (+), staphylokinase (+) and β-lysin(-), with repression of β-lysin caused by insertional inactivation of chromosomal based genes (Coleman *et al.*, 1989).

The environment in which bacteria exist can affect the transduction of bacteriophage and the acquisition of virulence factors, with *in vivo* factors playing a role in this (Mel and Mekalanos, 1996). Shiga toxin encoding bacteriophage have been found to be produced within the mammalian host intestine. They are capable of transducing naive *E. coli* strains within the gastrointestinal tract (Acheson *et al.*, 1998), the result of which is to amplify both toxin and bacteriophage production (Gamage *et al.*, 2003). Bacteriophage WO has been shown to transfer between strains of *Wolbachia* in the same cytoplasmic environment in the arthropod host (Masui *et al.*, 2000; Bordenstein and Wernegreen, 2004). While most examples are of lateral transfer occurring in the digestive tract, bacteriophage have been found to penetrate into many other environments in the mammalian host. Orally applied bacteriophage have been identified in the blood and urinary tract of both humans and mice (Dabrowska *et*

al., 2005). Intraperitoneal injection has also been an effective route for dissemination in the bloodstream (Merril *et al.*, 1996) although without sensitive bacteria to replicate the bacteriophage, the immune system can effectively remove bacteriophage (Dabrowska *et al.*, 2005).

While addition of MAV1 phage to *Mycoplasma arthritidis* upregulates virulence, one of the two transcribed bacteriophage genes, *vir*, instead codes for a lipoprotein located on the cell membrane. This is believed to exclude superinfecting phage at the level of DNA penetration (Clapper *et al.*, 2004). Temperate bacteriophage Sf6, which infects *Shigella flexneri*, encodes a gene *oac* which corresponds to O-antigen acetylase and causes a serotype conversion which results in the acquisition of immunity to superinfection with Sf6 (Clark *et al.*, 1991). Genes which encode for such external factors as cell membrane proteins, whether virulence factors or not, may be the basis for changes in colonial morphology noted in several cases (Barksdale and Arden, 1974). Infection with *tox*+ bacteriophage causes colonies of *Clostridium novyi* to become filamentous and spread in a snowflake like appearance (Eklund *et al.*, 1976). Infection with pf4 in *P. aeruginosa* biofilms results in small colony variants with increased surface attachment, possibly due to filamentous phage particles on the surface of the cells (Webb *et al.*, 2004). Similar small colony variants have been found with *E. coli* (Kuo *et al.*, 2000) although in this case, the size of colony was considered to be attributed to lysis and the presence of non-dividing cells.

2.8 Uses of Bacteriophage

2.8.1 Industrial uses

Bacteriophage of lactic acid bacteria are a serious problem in the dairy industry as they occur naturally in milk and can survive the pasteurisation process (Madera *et al.*, 2004). In addition, many starter cultures are lysogenic (Kilic *et al.*, 1996). This is of relevance as lactic acid bacteria are required for the fermentation of milk, for production of products such as yoghurt and cheese (Kilic *et al.*, 1996). Lysis of these bacteria can seriously damage the fermentation process. Control of environmental conditions has been shown to affect the amount of lysis in cultures containing

lysogens, thus affecting the bacterial load and properties of the dairy product (Lunde *et al.*, 2005). Lactic acid bacteria are also required in the fermentation of sauerkraut. In this situation a succession of various species at certain times are necessary for proper fermentation. Examination of the process has determined that bacteriophage may play an important role in this succession (Lu *et al.*, 2003).

The use of bacteriophage alone or in combination with antibiotics as protection from disease is also being examined in animal production with work underway with chickens (Huff *et al.*, 2004) and cattle (Smith *et al.*, 1987b). A cocktail of three lytic bacteriophage has been used experimentally to eliminate *E. coli* O157:H7 from beef carcases during slaughter (O'Flynn *et al.*, 2004).

Endolysin, a protein expressed by bacteriophage to lyse bacterial cell walls, has been recombinantly expressed and used experimentally to kill *Erwinia amylovora* (Kim *et al.*, 2004). This organism is the cause of fire blight, a disease causing economic loss in fruit orchards around the world.

2.8.2 Phage therapy

Bacteriophage can also be used to treat bacterial infections. Severe diarrhoea in calves has been cured with a single dose of bacteriophage. Disease prevention was possible by spraying litter with bacteriophage, or keeping calves in uncleaned rooms, previously housing animals treated with bacteriophage (Smith *et al.*, 1987b). Treatment with bacteriophage in combination with antibiotics has been found to completely eliminate mortality in broiler chickens (Huff *et al.*, 2004). An *in vitro* model of the human intestinal tract has been used to show that bacteriophage treatment of human intestinal disorders could also be effective (Verthe *et al.*, 2004). Recently a preparation of biodegradable polymer impregnated with ciprofloxacin and a mix of bacteriophage was used to successfully treat three men who had ulcerated wounds caused by radiation poisoning. The wounds were infected with antibiotic resistant *Staphylococcus aureus* and wound healing prior to this had not succeeded after one month of antibiotic treatment. After seven days exposure to the polymer, *S. aureus* was eliminated from the wounds, permitting healing to occur (Jikia *et al.*,

2005). The Institute of Immunology and Experimental Therapy in Poland claims a success rate of approximately 80% (2000 patients) in treating antibiotic resistant bacterial infections (Gorski and Weber-Dabrowska, 2005). Orally administered bacteriophage has also been shown to reach the peripheral blood and migrate to sites of infection (Dabrowska *et al.*, 2005). The oral environment can also be treated with bacteriophage. *Enterobacter faecalis*, an infection issue sometimes encountered in orthodontic treatment, was successfully eliminated using bacteriophage in an *in vitro* model (Paisano *et al.*, 2004). Furthermore, the use of bacteriophage encoded proteins for treatment of the oral environment has also been proposed and investigated (Nelson *et al.*, 2001; Hitch *et al.*, 2004).

2.8.3 Diagnostic uses

Bacteriophage can be used to diagnose disease by several methods. Bacteriophage capable of infecting and lysing the bacterium of interest are introduced to the suspected organism. The presence of the bacterium of interest is identified by either amplification and lysis resulting in plaques, or detection of a signal in the bacterium where the bacteriophage has been modified to produce one, for example by carrying a luciferase gene. An analysis of bacteriophage tests for the detection of *Mycobacterium tuberculosis* determined that they were currently highly specific but sensitivity was variable and performance was similar to current sputum tests (Kalantri *et al.*, 2005). As such, further work would be needed to replace current testing methods. Radio-labelled bacteriophage have been suggested for use in detecting inflammation caused by bacteria as opposed to sterile inflammation (Rusckowski *et al.*, 2004). As lytic bacteriophage have shown no evidence of side effects in humans, but are rapidly removed from the body by the immune system (Ochs *et al.*, 1971), bacteriophage can be used to examine the effectiveness of the immune system. After infection, the presence of bacteriophage and anti-bacteriophage antibodies is monitored to determine the responsiveness of the immune system. The bacteriophage ΦX174 has been used in this manner with several immunodeficiency diseases including Acquired Immunodeficiency Syndrome (AIDS) (Ochs *et al.*, 1971; Fogelman *et al.*, 2000).

2.8.4 Research

Bacteriophage have been central to discovery in molecular biology as mentioned previously (Section 2.1) and have been the source of useful molecular tools such as T4 DNA ligase (Lehman, 1974; Bernstein, 1981; Karam, 2005). They are also used to display proteins on their surface, a technique known as phage display (Russel *et al.*, 2004). In this process, filamentous phage are modified to display peptides or proteins on the capsid proteins. Proteins that have been displayed include antibody fragments (Hoogenboom *et al.*, 1991), growth factors (Dubaquie and Lowman, 1999), zinc finger proteins (Rebar and Pabo, 1994) and cocaine receptors (Carrera *et al.*, 2004), to name a few. The development of libraries displaying proteins allows for the identification and selection of genes encoding proteins with binding activities of interest (Russel *et al.*, 2004) and also potentially for use in medical treatment (Carrera *et al.*, 2004).

Bacteriophage coat proteins have been used to package RNA and DNA as positive controls for use in molecular detection of viruses (Walkerpeach and Pasloske, 2004). This had the advantage of stabilising the nucleic acid against degradation by ribonucleases (RNases) and deoxyribonucleases (DNases). Other research has included examination of outer membrane proteins of bacteria in real time as the bacteria grow and divide. In this case labelled bacteriophage lambda tails, which bind to these proteins, were used (Gibbs *et al.*, 2004). Identification of bacteriophage of actinomycetes in soil samples has been used to indicate the presence of slow growing actinomycetes in the soil (Gathogo *et al.*, 2004).

2.9 *Burkholderia pseudomallei*

Burkholderia pseudomallei is a Gram-negative saprophytic rod and a facultative intracellular pathogen (Jones *et al.*, 1996). It is the causative agent of the disease melioidosis which occurs primarily in tropical and subtropical regions (Dance, 2000). Melioidosis has been studied for weaponisation in several countries and *B. pseudomallei* is considered a potential bioterrorism tool (Bossi *et al.*, 2004; Cheng *et al.*, 2005a).

2.9.1 Initial identification

Human melioidosis was first documented in Rangoon, Burma by Captain Alfred Whitmore and C.S. Krishnaswami (Whitmore, 1913) from 38 fatal cases, beginning in 1911. Symptoms similar to those of glanders, caused by *Bacillus mallei*, were identified, so Captain Whitmore named the pathogen *Bacillus pseudomallei.*

2.9.2 Taxonomy of *B. pseudomallei*

The initial nomenclature of the organism now known as *B. pseudomallei* was, as noted above, *Bacillus pseudomallei.* Further nutritional, biochemical and morphological studies have resulted in name changes to *Bacillus whitmori, Malleomyces pseudomallei, Loefflerella whitmori, Pfiefferella whitmori* and *Pseudomonas pseudomallei* (Cottew *et al.*, 1952; Kingston, 1971; Smith *et al.*, 1987a; Dance, 1991). The current nomenclature was based on genetic studies (Lew and Desmarchelier, 1993; Gillis *et al.*, 1995).

2.9.3 Geography

Since 1911, *B. pseudomallei* has primarily been identified in countries between latitudes 20° north and 20° south (Dance, 1991). It has been identified as a common cause of community-acquired septicaemia and death in northeastern Thailand (Chaowagul *et al.*, 1989) with a mortality of 68%. In Australia, *B. pseudomallei* infection was first identified in sheep in 1949 near Winton, Queensland (Cottew *et al.*, 1952), with the first human case of melioidosis being identified in 1950 in Townsville, Queensland (Rimington, 1962). The overall mortality for melioidosis in the Northern Territory of Australia has been calculated at 21% (Currie *et al.*, 2000). Cases of melioidosis in countries outside the 20° north and south latitudes have also been identified and attributed to importation of infected animals (Currie *et al.*, 1994) and humans (Dance *et al.*, 1999; Peltroche-Llacsahuanga and Haase, 1999; Carlson and Seppanen, 2000; Karcher *et al.*, 2000; Shrestha *et al.*, 2005).

2.9.4 Epidemiology

Infection with *B. pseudomallei* is more common in those who have a close association with the soil and water, such as rice farmers in Thailand (Chaowagul *et*

al., 1989) and during the monsoon season in Australia (Currie *et al.*, 2000). The route of infection is inoculation via injuries to the skin or by inhalation or ingestion (Barnes and Ketheesan, 2005). The presence of immunosuppressant factors such as diabetes, renal disorders (Chaowagul *et al.*, 1989), excessive consumption of alcohol, chronic lung disease and malignancy or corticosteroid use (Currie *et al.*, 2000) have been associated with melioidosis. Australian aborigines in endemic regions have a higher rate of development of melioidosis than the rest of the community (Currie *et al.*, 2004) probably due to a lifestyle with frequent exposure to soil and surface water. In Australia, human serological studies have identified an antibody response to *B. pseudomallei* ranging from 5.7% in northern Queensland to 12.8% in the Northern Territory (Currie *et al.*, 2000).

2.10 Clinical Presentation of Melioidosis

Symptoms of infection with *B. pseudomallei* can range from an asymptomatic carrier state to septicaemia resulting in death within 48 hours (White, 2003). Because of the wide range of presentations, it is known as the great mimicker of disease (Yee *et al.*, 1988) and can be confused with several other disorders. Most organ systems can be affected (Table 2.2) and melioidosis is often categorised into acute, subacute and chronic forms with the possibility of conversion between these forms (Kingston, 1971). Recrudescence of the disease has been reported months to years after the initial infection (Desmarchelier *et al.*, 1993; Leelarasamee, 1998; Ngauy *et al.*, 2005) as has latent infection converting to active forms of the disease (Mackowiak and Smith, 1978).

2.11 Humoral and Cellular Responses to *B. pseudomallei*

Exposure to *B. pseudomallei* may lead to the formation of specific antibodies and the development of cell-mediated, adaptive immunity in individuals who do not develop symptoms of melioidosis (Govan and Ketheesan, 2004). Those who survive melioidosis have also been shown to have developed a cell mediated, adaptive immune response (Ketheesan *et al.*, 2002).

Table 2.2 Common clinical presentations of melioidosis (from; Leelarasamee and Bovornkitti, 1989)

2.12 Virulence Factors of *B. pseudomallei*

Virulence factors of *B. pseudomallei* include a type three secretion system (TTSS) (Winstanley *et al.*, 1999), an exotoxin (Mohamed *et al.*, 1989a) and flagella (Chua *et al.*, 2003), with potential virulence determinants including surface polysaccharides, fimbrae, pili, drug resistance determinants and putative adhesins (Holden *et al.*, 2004).

Pathogenicity islands (PIs) are virulence genes found in large contiguous groups and TTSSs have been associated with these. TTSSs are made up of a number of proteins homologous with components of flagellum-specific export apparatus and are involved in delivering virulence factors directly to host cells, with secretion being triggered by a pathogen coming in close contact with host cells (Mecsas and Strauss,

1996). A TTSS gene cluster has been identified on *B. pseudomallei* (Winstanley *et al.*, 1999) and been found to be typically absent in the avirulent variant *B. thailandensis* (Winstanley and Hart, 2000).

Conserved TTSSs are common in bacterial pathogens of plants and animals (Keen *et al.*, 2000). Examples of bacteria using TTSSs include *Yersinia enterocolitica* which contains a plasmid encoded TTSS (Michiels and Cornelis, 1991), *P. aeruginosa* which translocates several toxins including an ADP- ribosylating exotoxin (*ExoS*) (Yahr *et al.*, 1996; Rietsch *et al.*, 2004) and *Salmonella enterica* which encodes two separate TTSSs (SPI1 and SPI2). SopE is one of the proteins translocated by SPI1 and is found on a prophage distant from the TTSS PI, as are several other effector proteins. It has also been suggested that the TTSS was itself acquired through horizontal gene transfer at some previous time (Mecsas and Strauss, 1996; Yahr *et al.*, 1996; Hardt *et al.*, 1998; Mirold *et al.*, 1999; Mirold *et al.*, 2001). It is possible that some of the virulence factors translocated by the TTSS of *B. pseudomallei* could also be bacteriophage encoded.

An exotoxin of *B. pseudomallei* has been identified to be ADP-ribosylating (Mohamed *et al.*, 1989a; Mohamed *et al.*, 1989b). ADP-ribosylating toxins in some other bacteria such as the diphtheria toxin in *Corynebacterium diphtheriae,* mentioned previously (Freeman, 1951; Collier and Kandel, 1971; Laird and Groman, 1976a; Laird and Groman, 1976b), the C3 toxin of *Clostridium botulinum* (Eklund *et al.*, 1971; Popoff *et al.*, 1990) and the cholera toxin of *Vibrio cholerae* (Trepel *et al.*, 1977; Waldor and Mekalanos, 1996) are known to be bacteriophage mediated. Bacteriophage T4 (family *Myoviridae*) has also been shown to code three ADP-ribosylating activities (Wilkens *et al.*, 1997; Depping *et al.*, 2005). It would not be unusual for the ADP-ribosylating exotoxin of *B. pseudomallei* to also be bacteriophage encoded.

In some bacteria, cell surface proteins used for adhesion to host cells have been identified as being similar to bacteriophage coat and capsid proteins. In *V. cholerae,* it has been reported that the toxin co-regulated pilus which plays a role in host

colonisation is encoded on the putative cryptic bacteriophage VPIΦ (Karaolis *et al.*, 1999). In *Streptococcus mitis*, platelet binding is promoted by two large cell surface proteins encoded on bacteriophage SM1 (Bensing *et al.*, 2001). As noted previously, there are several cell surface related potential virulence determinants present in *B. pseudomallei*, some of which may be bacteriophage encoded.

2.13 An Isolate of Interest

The NAFC isolate used in this study was developed by serial plating NCTC 13178 on nutrient agar containing 0.1% ferric citrate in a successful attempt to attenuate virulence (Ulett *et al.*, 2001). There was also a distinct alteration in colony morphology coinciding with attenuation in virulence and an increase in agar pH was observed. Ulett *et al.* (2001) concluded that the attenuation of virulence was dependant on iron availability, although other isolates tested using this method did not result in attenuation or changes in colony morphology. Also, attempts to upregulate virulence by removal of ferric citrate and by passage through mice were unsuccessful. Haemolysin activity, a known virulence factor which can be regulated by iron (Dai *et al.*, 1992), was shown to be reduced on sheep blood agar (SBA) in the case of NAFC. There was no discussion by Ulett *et al.* (2001) of why the changes that produced NAFC were not reversible as would be expected when changes were initiated by regulation of expression. One hypothesis would be that mutations to a bacteriophage have caused these changes and addition of unmutated bacteriophage to the bacteria could restore lost function.

It is possible that during the serial passage on iron rich medium, the removal of selective pressure to maintain iron regulation meant mutations (for iron regulation) to the genome were not selected out. These alterations may have been due to one or more base mutations (Lawrence and Roth, 1996), resulting in either loss of a regulatory or structural gene or group of genes. It should also be noted that some bacteria contain prophage deletion processes that clear bacteria of non-essential DNA (Boyd and Brussow, 2002). Such changes cannot be reversed merely because selective pressure has recommenced. Another possibility is that a second bacteriophage has infected the bacteria and has entered at a site that disrupts iron

regulation similar to the way serotype F bacteriophage disrupts β-lysin production in *S. aureus* (Coleman *et al.*, 1989).

Also of note, was the striking alteration of colony morphology. This could also be caused by mutations which affect colonial morphology. As has been noted previously (Eklund *et al.*, 1976; Kuo *et al.*, 2000; Webb *et al.*, 2004), bacteriophage can affect colony morphology of its host and changes in the bacteriophage could hence cause further alteration or a loss of the original alteration.

2.14 Bacteriophage Infection of *B. pseudomallei*

In 1956, Leclerc and Sureau collected bacteriophage from the waters around Saigon and Hanoi that were capable of lysing *B. pseudomallei*, indicating the presence of *B. pseudomallei* in the waters (Leclerc and Sureau, 1956). *B. pseudomallei* isolates were found to be variably sensitive to the bacteriophage, which specifically lysed *B. pseudomallei.* In later work (Denisov and Kapliev, 1991; Denisov and Kapliev, 1995), bacteriophage were extracted from *B. pseudomallei* with the aid of chloroform. In these cases, *B. pseudomallei* isolates were also found to be variably sensitive to the bacteriophage extracted. Most *B. pseudomallei* isolates examined were found to be lysogens and some were found to be polylysogenic. That is, they produced more than one bacteriophage (Denisov and Kapliev, 1991). The bacteriophage extracted from *B. pseudomallei* were found to belong to the *Myoviridae* and *Syphoviridae* families (Denisov and Kapliev, 1995). Family *Syphoviridae* bacteriophage have also been identified in the closely related bacterium *Burkholderia thailandensis* (Woods *et al.*, 2002). Examination of *B. pseudomallei* genomes by subtractive hybridisation (DeShazer, 2004) has identified the presence of a prophage in *B. pseudomallei* isolate 1026b, which could also be spontaneously induced and was of the family *Syphoviridae*. Complete sequencing of the genome of *B. pseudomallei* K96243 (Holden *et al.*, 2004) identified genomic islands indicative of horizontal transfer, at least three of which appeared to be prophages. One of these (ΦK96243; also known as genomic island two) was successfully induced and found to belong to the family *Myoviridae*.

Smith and Cherry (1957) determined that *Burkholderia mallei* is sensitive to *B. pseudomallei* bacteriophage and this has been used by others as a host for *B. pseudomallei* bacteriophage (Denisov and Kapliev, 1991; DeShazer, 2004) and for identification of *B. mallei* using bacteriophage (Woods *et al.*, 2002). The use of *B. mallei* in this fashion has risks as it has resulted in laboratory acquired infections (Srinivasan *et al.*, 2001) producing the life threatening disease glanders.

2.15 Conclusions

In summary, bacteriophage are widespread in many environments on Earth and play roles in bacterial ecologies and pathogenesis. They have been used in developing an understanding of molecular biology, as molecular tools and to treat bacterial diseases. The human pathogen *B. pseudomallei* carries virulence factors which in some other organisms are encoded on bacteriophage. Isolates of *B. pseudomallei* have been identified as lysogens and bacteriophage may also play a role in virulence in this organism.

CHAPTER 3

GENERAL MATERIALS AND METHODS

3.1 Bacteria

3.1.1 Isolate origins and identification

Bacteria used in this work were sourced from the James Cook University (JCU) Microbiology bacterial collection and from samples collected in Papua New Guinea by Dr Jeffrey Warner from Microbiology and Immunology, JCU. The JCU Microbiology collection includes isolates collected from local soil, the Department of Pathology, Townsville General Hospital, and The Menzies School of Health Research, Royal Darwin Hospital, Darwin. Identity of the isolates was previously confirmed by colonial morphology on Ashdown agar and API20NE (bioMerieux, La Balme, France).

3.1.2 Stock suspensions

Stock suspensions of *B. pseudomallei* were prepared by inoculating single colonies from Ashdown agar (appendix 1) into ten ml of brain-heart infusion broth (BHIB) (Appendix 1) and incubating at 37°C for 18 hours in an orbital shaker (100rpm; Orbital Mixer Incubator, Ratek, Australia). Following incubation, the suspension was made to 10% glycerol (v/v) and one ml aliquots were frozen at -70°C until used. When required, aliquots were thawed at room temperature.

3.1.3 Quantitation of bacteria

Bacterial concentration in suspension was determined using a variation on a Miles-Misra assay (Miles *et al.*, 1938; Slack and Wheldon, 1978) in which serial ten fold dilutions of bacterial suspensions in phosphate buffered saline (PBS) (Appendix 1) were carried out and 20µl of each dilution aliquoted and spread onto sheep blood agar (SBA) in triplicate (Appendix 1). The plates were allowed to dry and then incubated at 37°C for 24-48 hours and colony number in the lowest measurable dilution recorded. The following equation was used to determine bacterial concentration.

cfu/ml = average colony count × 50 × reciprocal of dilution

for example; an average colony count of 24 colonies at a 10^{-4} dilution would be 1.2×10^7 cfu/ml = $24 \times 50 \times 10^4$

3.1.4 Production of competent cells

Single colonies of JM109 (Promega, USA) or XL1-Blue MRF (Stratagene, USA) *E. coli* from an overnight incubation on Luria-Bertani (LB) agar were inoculated in 2.5ml LB broth and incubated overnight at 37°C with shaking at 130rpm (Sepatech suprafuge, Heraeus, Germany). The broth was then added to 250ml LB in a one litre conical flask and incubated for one to five hours at 37°C with shaking at 130rpm or until the optical density at 600nm ($O.D._{600nm}$) was between 0.4 and 0.6. The broth was centrifuged at 4°C for ten minutes at 4500*g*. The supernatant was discarded and the pellet resuspended in 2.5ml of ice cold 0.1M calcium chloride ($CaCl_2$) (Appendix 1). This was chilled on ice for 30 minutes and the solution centrifuged at 4°C for five minutes at 4000*g*. The supernatant was discarded and the pellet resuspended again in 2.5ml of ice cold 0.1M $CaCl_2$. To this, 375µl of 50% glycerol was added and the solution stirred gently with a pipette prior to dispensing into 200µl aliquots in screw capped microfuge tubes with o-rings (Quantum Scientific, Australia). These were then snap frozen in liquid nitrogen and stored at -80°C until required. Typically cells produced by this method had a cloning efficiency of 10^5 cfu/µg of DNA.

3.2 DNA Preparation

3.2.1 Genomic DNA extraction

All genomic DNA extractions were from fresh overnight broths of bacteria. Genomic DNA was extracted using either a commercial kit (Genomic Tip 100/G, Qiagen), following the kit protocol or extracted based on the method of purification of DNA from aqueous solutions in Short Protocols in Molecular Biology by Ausubel (1992). In detail the latter method is as follows (all media preparations are detailed in Appendix 1):

The bacteria of interest was streaked on Ashdown agar and incubated at 37°C for two days. Five ml of tryptone soya broth (TSB) was inoculated with a single colony and

incubated with shaking for two days (37°C, 100rpm). This culture was aliquoted (1.5ml) into microfuge tubes (Sarstedt; Germany) and pelletted (13000*g*, two minutes, Universal 16R centrifuge, Hettich, Germany). The supernatant was removed and each pellet resuspended in 567µl of Tris-EDTA (TE) buffer. Thirty µl of 10% sodium dodecyl sulphate (SDS) and three µl of 20mg/ml proteinase K (Sigma) were added and the solution mixed and then incubated (37°C/one hour). At this time 100µl of five molar sodium chloride (NaCl) and 80µl CTAB/NaCl solution were added. The solution was mixed and incubated for ten minutes at 65°C and then cooled to room temperature. Ribonuclease A was then added to a final concentration of ten µg/ml and the solution incubated for one hour. After incubation, an equal volume of chloroform/isoamyl alcohol (24:1) (Sigma) was added and the mixture vortexed for 30 seconds, followed by centrifugation for five minutes. The upper aqueous phase was removed and transferred to a fresh tube with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) added. This was vortexed and centrifuged as above and the supernatant removed and transferred to a fresh tube. Isopropanol was added (0.6 volumes based on supernatant volume) and the solution mixed gently before freezing at -70°C for one hour to facilitate precipitation. The solution was thawed, centrifuged (10 000*g* / five minutes) and the supernatant removed. The pellet was washed with 70% ethanol (Sigma, Australia) and then with 95% ethanol. The pellet was dried at 37°C and resuspended in 100µl of TE buffer.

3.2.2 Ethanol precipitation

An alternative to the addition of isopropanol for precipitation of DNA is that of ethanol precipitation. This was used when bacteriophage DNA was extracted and can be described as follows;

One tenth of a volume of sodium acetate (10% solution, pH5.2) was added to the aqueous solution of DNA (one volume - typically up to 700µl) and the solution was inverted to mix. To precipitate the DNA 1.1 volumes of 95% ethanol was added. The DNA was pelleted at 20 000*g* for 30 minutes at 4°C (Sepatech suprafuge, Heraeus, Germany) and then washed with 70% ethanol and pelleted again (15 minutes, 20 000*g*, 4°C). The ethanol was removed and the pellet dried for 30

minutes in a vacuum drier at 32°C (Centrivap Concentrator; Labconco). This dried DNA sample was made up in an appropriate buffer such as sterile distilled water, Tris buffer or TE buffer and stored at -20°C until required.

3.2.3 Ligation and transformation

3.2.3.1 Preparation of vectors

Plasmid DNA was digested with the desired restriction endonuclease (Promega) according to manufacturer's instructions with the exception that the digestion was scaled up to ten µg of DNA in a total of 100µl of digestion mix and digestion was carried out for 1.5 hours. For each ten µg of plasmid DNA digested in a total volume of 100µl, ten µl of 200mM ethylenediaminetetraacetic acid (EDTA) was added to stop the restriction endonuclease reaction. Then 20µl of ten mM tris.HCl pH8.3, 15µl of 10x calf intestinal alkaline phosphatase (CIAP) buffer (Promega) and 6.8µl of 0.01Units/µl of CIAP (Promega) was added. This amount of CIAP equates to sufficient for 6.8 pmol of ends, with the pmol of ends required calculated as shown in the following equation.

$$\text{pmol required} = \frac{\text{µg of plasmid} \times 3.03}{\text{size of plasmid in kb}}$$

The reaction was digested at 37°C for 30 minutes and a second 6.8µl aliquot of CIAP was added. The sample was digested for another 30 minutes at 37°C and immediately cleaned up using a DNA clean up kit (Wizard® SV Gel and PCR Clean-Up System) with elution into 100µl of ten mM Tris.HCl or deionised water. $O.D._{260/280nm}$ was determined as per Section 3.2.5 and the sample was stored at -20°C until required. Typically 50% of DNA was lost during this method.

3.2.3.2 Ligation

Ligations of PCR products were carried out using either the pGEM®-T Vector System (Promega) or T&A Cloning Vector Kit (RBC, Taiwan) according to the manufacturer's instructions. Ligations of cut DNA with restriction endonuclease were carried out as described below;

The amount, in nanograms, of the insert to be added to 50 or 100ng of restriction cut, dephosphorylated plasmid was calculated as follows

$$\frac{\text{(ng vector)} \times \text{(kb insert)} \times \text{(insert:vector ratio)}}{\text{(kb vector)}} = \text{ng insert}$$

e.g. When using 100ng of a 4.5kb plasmid and a 0.7kb insert, with a 3:1 insert to vector ratio

$$\frac{100 \times 0.7 \times 3}{4.5} = 47\text{ng of insert}$$

Ligations were carried out overnight at 4°C in a total volume of 10µl. T4 DNA ligase buffer and T4 DNA ligase were sourced from Promega. Controls were included to test the quality of the reagents. An example of ligation ingredients is shown in Table 3.1.

Table 3.1 Example of ligation experiment, including sample and two controls. The Dephosphorylation control tests the efficacy of the dephosphorylation step. The insert control tests that the insert will not self ligate and be transformed. The total volume of each reaction is ten µl

	Sample 3:1 insert to vector ratio (µl)	Control: Dephosphorylation (µl)	Control: insert (µl)
plasmid (100ng/µl)	1	1	
insert (700bp @ 47ng/µl)	1		1
T4 DNA ligase (3 Weiss units/µl)	1	1	1
10x Ligase buffer	1	1	1
dd H_2O	6	7	7

3.2.3.3 Transformation

After overnight incubation, transformation was carried out with either commercial competent cells using the manufacturer's instruction or with in house produced competent cells (Section 3.1.4). In the case of in house competent cell use, 2ul of

ligation mix was added to chilled 500ul thin walled microfuge tubes on ice. Competent cells were thawed on ice for about 5 minutes and flicked with a finger to mix. One hundred microlitres of cells were added to the chilled ligation mix (two µl) using a chilled pipette tip, this was flicked to mix, then stored on ice for 20 minutes. The tubes were heat shocked for 45-50 seconds at 42ºC , then stored on ice for two minutes. Three hundred µl of room temperature SOC medium (Appendix 1) was added and the total volume removed and placed in a 1.5ml microfuge tube with 600µl additional SOC medium. This was incubated for 1.5hrs at 37ºC, with shaking at 100rpm. One hundred microlitres of the broth was then plated on LB agar plates containing 80µg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) (Astral; Australia), appropriate antibiotics, according to the plasmid used, and surface plated with 20µl of 500mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Astral, Australia). These were incubated at 37°C overnight then placed at 4°C for one hour to enhance any blue colouration. White colonies were selected for the purpose of growth in LB broth with appropriate antibiotics present and plasmid extraction.

3.2.4 Plasmid DNA extraction

Plasmids were extracted from fresh overnight broths of bacteria in LB medium (Appendix 1) containing the relevant antibiotic at the appropriate concentration for the plasmid in question (Table 3.2). Plasmid DNA was extracted using one of the following commercial kits; Wizard® *Plus* SV Minipreps DNA Purification System (Promega), FastPlasmid™ Mini (Eppendorf, Germany) or HiYield™ Plasmid Mini Kit (RBC, Taiwan), in accordance with the manufacturer's instructions.

Table 3.2 Plasmids used and antibiotic added to agar plates and liquid media

Plasmid	Antibiotic
pBK-CMV (Stratagene, USA)	kanamycin (LB with 0.5% NaCl - 25µg/ml LB with 1.0% NaCl - 50µg/ml)
pGEM-T (Promega, USA)	ampicillin (100µg/ml)
T&A cloning vector (RBC, Taiwan)	ampicillin (50µg/ml)

3.2.5 Determination of DNA concentration

The concentration of DNA was determined by absorbance of light at 260nm ($O.D._{260nm}$), using a UV mini 1240 spectrophotometer (Shimadzu, Japan) or BioPhotometer (Eppendorf, Germany), where one unit of absorbance is equivalent to a double stranded DNA (dsDNA) concentration of 50µg/ml. The purity of the dsDNA was determined by comparison of the $O.D._{260nm}$ and $O.D._{280nm}$, where a result of 1.8 for $O.D._{260nm}/O.D._{280nm}$ is indicative of pure DNA. Lower readings indicate protein contamination and higher readings indicate RNA contamination (Sambrook and Russell, 2001).

3.2.6 Polymerase chain reaction (PCR)

3.2.6.1 General apparatus

All PCR reactions were carried out in either a PTC-100™ (MJ Research Inc., Massachusetts, USA) or a Mastercycler® gradient (Eppendorf, Hamburg, Germany) thermocycler using 200µl thin walled PCR reaction tubes (Pathtec, Australia). General reagents used included molecular biology grade water (Sigma, Australia), 5U/µl *Taq* DNA polymerase (recombinant), 10x PCR Buffer with ammonium sulphate ($(NH_4)_2SO_4$) and 25mM magnesium chloride ($MgCl_2$) (MBI Fermentas, Germany), 10mM PCR nucleotide mix (dNTP) (Promega) and 0.05µmole lyophilised custom primers (Sigma-Genosys, Australia). Custom primers were made up to 100pmol/µl (100µM) using the protocol described in Section 3.2.8.

3.2.6.2 PCR reagents

Initial PCR reactions were carried out using a standard set of reagent concentrations with optimisation of each reaction being done separately. A master mix was prepared using reagents in the concentrations shown in Table 3.3, to a final volume of 50µl per tube.

Table 3.3 General reagents used in PCR reactions

Reagent	Concentration or Volume
molecular biology grade water	to 50µl
Reaction buffer (10x)	5µl
$MgCl_2$ (25mM)	variable (generally 2-3mM)
dNTP (10mM)	200µM each
Primers (100pmol/µl)	1µM each
Recombinant Taq DNA polymerase (5U/µl)	1.5µl
Template DNA	10-30ng

3.2.6.3 Cycling parameters

The cycling parameters used for each PCR reaction varied depending on the melting temperature (Tm) of the primers being used and the length of the expected amplicon. Listed below is a general overview with specific sets of cycling conditions given in detail in appropriate chapters.

- Denaturation at 94°C for ten minutes
- 36 cycles of;
  1. denaturing at 94°C for one minute
  2. annealing of primers at 55 – 60°C for 45-60 seconds
  3. extension at 72°C for 1-3.5 minutes (~1 minute/1000bp of product)
- extension at 72°C for four minutes
- storage at 15°C in the thermocycler until removed and stored at 4°C prior to further processing.

3.2.7 Agarose gel electrophoresis

PCR products were visualised using agarose gel containing 0.5µg/ml ethidium bromide. Gels were between 0.8% and 2% agarose (Progen molecular biology grade agarose) (W/V) in Tris acetate (TAE) buffer (Appendix 1) depending on the size of the DNA product. Smaller products (<1000bp) were run on 2% gels of 100ml volume (Liberty 1 tank, Fisher Biotec, Australia) at 200V. Larger products were visualised using 0.8% agarose gels of 150ml volume (HU13 Tank, Scie-Plas, Crown Scientific, Australia) at 90V.

Ten µl of PCR product was added to two µl of 6x loading dye (Appendix 1) and placed into preformed wells in the agarose gels.

Fluorescent signals from ethidium bromide intercalated in DNA bands under UV light were visualised and captured using GelDoc 1000 (Biorad, Australia) and GeneSnap v.4.00.00 (Syngene, Synoptics, England) respectively. The size of DNA products was estimated empirically by comparison of migration distances from the loading well of DNA and molecular weight marker (GeneRuler™ 1kb DNA ladder, Fermentas, Germany, 1kb DNA Ladder, Promega, Australia or 1kb Plus DNA Ladder™ , Life Technologies Australia) using GeneTools v.3.00.22 (Syngene, Synoptics, England).

3.2.8 Primer design

Primers were designed using the program Oligo v.6.65 (Molecular Biology Insights Inc., USA). Sequence data was loaded into the program and primers were designed under high stringency with a length of 18-21 bases. Primers were selected in accordance with the noted criteria:

1. False priming site efficiency must be less than 25% of correct priming site efficiency.
2. There must be no hairpins consisting of three or more base pair stems.
3. Primers must not form strong or 3' terminal duplexes with themselves or each other.

If primers were to be used for sequencing, sequencing primers rather than compatible pairs were selected in the Oligo "search for primers and probes" selection screen. All primers (Genosis, Sigma, Australia) were supplied lyophilised in quantities of at least 0.05µmole. They were reconstituted to a final concentration of 100pmol/µl using molecular biology grade water before storage at -20°C. The reconstitution protocol is described below;

The tube containing lyophilised primer was briefly centrifuged by pulsing. The volume of water added was ten fold the nmols of DNA present in the tube. For example, 125nmol would have 1250µl of water added. The nmol information was provided on the Sigma data sheet. The water was added and the tube inverted several

times. The tube was allowed to stand for two minutes before being pulsed down briefly. The inversion, standing and pulsing down was repeated twice.

3.2.9 Sequencing

All sequencing of plasmid inserts was carried out at either James Cook University (see below), or Macrogen (Korea). Samples sent to Macrogen were in the form of plasmids with inserts at a concentration of 50ng/µl or more. Sequencing at JCU was done as follows; sequencing PCRs were carried out in a Mastercycler® thermocycler using 200µl thin walled PCR reaction tubes containing 20µl of of reagents (Table 3.4). The cycling parameters were 34 cycles of; 95°C for 20 seconds, 50°C for 15 seconds, 60°C for one minute and 20 seconds, followed by storage at 10°C in the thermocycler until removed and processed for sequencing.

All primers were synthesised by Sigma Aldrich, Australia. An Amersham DYEnamic ET Teminator Cycle Sequencing Kit (Amersham Biosciences, NSW, Australia) was used as described in Table 3.4. Post PCR clean-up was carried out using AutoSeq™ G-50 columns (Amersham Biosciences, NSW, Australia) according to the manufacturer's instructions. Reactions were analysed at the Genetic Analysis Facility (GAF) at James Cook University, Townsville, QLD, Australia, using a Beckman CEQ 2000 DNA Analysis System or an Amersham Biosciences MegaBASE 1000. Sequence data was cleaned and analysed using Sequencher™ (Gene Codes Corporation, Ann Arbor, MI, USA) and any vector or primer sequences were removed.

Table 3.4 Reagents used in sequencing PCR. Volumes are described for a total volume of 20µl. Reagents with an asterix were supplied with the DYEnamic ET Teminator Cycle Sequencing Kit. The primer used for the sample mix were for the primer sites present on the plasmid.

Reagent	Sample mix volumes or concentrations	Control mix volumes or concentrations
DYEnamic ET Teminator master mix	8µl*	8µl*
Primer	~5pmol	4µl control DNA*
DNA	400-600ng plasmid DNA	1µl M13mp8 control DNA*
water	to 20µl	7µl

3.3 Experimental Animals

Mice were purchased from the Small Animal Breeding Unit, School of Veterinary and Biomedical Sciences, JCU, Townsville. Experimental animals were aged 8-16 weeks and were housed in wire-topped, plastic boxes containing wood shavings as bedding. Males and females were used at a ratio as close to 1:1 as possible. Animals were maintained in an air-conditioned negative pressure isolation unit, fed protein enriched pellets and drinking water was provided *ad libitum*. Experimental animals were monitored daily and bedding was changed as required. Where possible, all experimental procedures were performed in a class II biological safety cabinet (Gelman Sciences, Australia). Bacterial manipulations were performed in the PC3 laboratory of JCU. Ethics approval for animal experimentation was granted under the approval numbers A526 and A978.

3.3.1 Determination of ID_{50}

Calculations of 50% endpoints of infectious dose (ID_{50}) of *B. pseudomallei* isolates in BALB/c mice were made using the procedure outlined by Ulett (1996) (Appendix 4) based on the earlier paper by Reed and Meunch (1938).

CHAPTER 4
IDENTIFICATION, EXTRACTION AND PROPAGATION OF BACTERIOPHAGE

4.1 Introduction

Many bacteria are, or can be, infected by bacteriophage. One study has estimated that up to 17% of all marine bacteria are infected by bacteriophage, whether lytic (virulent) or temperate (Suttle, 1994). Another more recent study showed over 40% of marine bacterial isolates carried inducible temperate bacteriophage (Jiang and Paul, 1998). As bacteriophage are common, they may be playing a large role in the survival of bacteria.

In many cases, the bacteriophage have been found to encode genes which give the bacteria a survival advantage or increased virulence (Uchida *et al.*, 1971; Scotland *et al.*, 1983; Smith *et al.*, 1983; Gertman *et al.*, 1986; Hayashi *et al.*, 1993; Hardt *et al.*, 1998; Mirold *et al.*, 1999; Mirold *et al.*, 2001; Strauch *et al.*, 2001; Williams *et al.*, 2002). Bacteriophage, or proteins produced by them, are being examined for their use in clinical treatment (Nelson *et al.*, 2001; Schuch *et al.*, 2002) or food processing (Shearman *et al.*, 1994; de Ruyter *et al.*, 1997; Yoon *et al.*, 2001; Randerson, 2003). Bacteriophage of the families *Myoviridae* and *Siphoviridae* have been previously identified in *B. pseudomallei* (Denisov and Kapliev, 1995) but, with the exception of a bacteriophage of the family *Siphoviridae* carried by the related bacterium *Burkholderia thailandensis* (Woods *et al.*, 2002), little more has been reported on the characterisation of bacteriophage from *B. pseudomallei.*

To examine bacteriophage from *B. pseudomallei* for both bactericidal and virulent properties, lysogens must first be identified, the bacteriophage extracted and concentrated or amplified as required, and the DNA extracted.

Traditional methods of screening and amplifying *B. pseudomallei* for bacteriophage (typically spontaneously produced bacteriophage) involve the use of *B. mallei* as a sensitive host (Denisov and Kapliev, 1991; Denisov and Kapliev, 1995; Woods *et al.*, 2002). *Burkholderia mallei* is highly infectious to humans, causing a life threatening

disease (Smith, 1957) and has been known to cause infections via laboratory contamination (Howe and Miller, 1947; Srinivasan *et al.*, 2001). As such, alternative methods of screening for bacteriophage and amplification are desirable.

The earliest identification of bacteriophage infecting *B. pseudomallei* involved the plating of water samples potentially carrying bacteriophage onto plates spread with a typical *B. pseudomallei* isolate. This protocol was carried out for several months on many water samples before plaques were noted. Numerous passages of the bacteriophage on this isolate of *B. pseudomallei* were required before sufficient amplification occurred to cause confluent lysis on a plate of *B. pseudomallei*. In addition, several isolated bacteriophage showed variation in plaque formation when tested against other isolates of *B. pseudomallei* (Leclerc and Sureau, 1956).

The basic method for detecting the presence of bacteriophage involves lysis of a host or sensitive strain of bacteria leading to a drop in optical density of a broth, or formation of plaques on a plate (Billing, 1969). For this to be observed in lysogenic systems, lysis may need to be triggered, usually by exposure to ultraviolet light or addition of mitomycin C. Mitomycin C has strong bactericidal effect against both Gram positive and Gram negative cells. It functions by inhibiting DNA synthesis by reacting covalently with DNA to form cross links between complementary strands (Ueda and Komano, 1984). This activates DNA repair mechanisms, including the RecA protein which cleaves a repressor and induces prophage (bacteriophage existing as DNA encoded within the bacterial genome) to enter the lytic cycle (Weinbauer and Suttle, 1996). Ultraviolet radiation at 254nm (UV-C) damages DNA. Similar to the mitomycin C assay, it triggers DNA repair mechanisms such as RecA (Kidambi *et al.*, 1996) and has previously been used to trigger a lytic bacteriophage cycle (Loessner *et al.*, 1991; Kidambi *et al.*, 1996; Weinbauer and Suttle, 1996; Woods *et al.*, 2002). In these studies, incubation times between exposure and harvesting of bacteriophage have varied from between 15 minutes to 50 hours (Jacob and Fuerst, 1958; Hooper *et al.*, 1981; Loessner *et al.*, 1991; Weinbauer and Suttle, 1996).

Lysogenic systems typically produce lower levels of bacteriophage than lytic systems and amplification of yield may be required to aid identification. While lysogenic cultures always contain some free bacteriophage produced by spontaneous lysis of an occasional cell, propagation of large amounts of bacteriophage requires induction under precise conditions, such as the addition of mitomycin C. If conditions of this assay are not ideal for the individual system, yield may not be optimal (Douglas, 1975). This, combined with the fact that mitomycin C is a highly toxic chemical, indicates that alternative means of producing bacteriophage are desirable. It has been noted that temperate bacteriophage can cause spontaneous plaque formation on host strains of bacteria. A system involving infection into a host in which the bacteriophage spontaneously enters the lytic lifecycle at a greater rate, will result in higher yields without the use of toxic chemicals.

There are several variations on the method for screening for the presence of bacteriophage. These include use of bacterial broth for amplification prior to plating, as noted above, and plating onto a soft agar overlay of host bacteria, or directly onto host bacteria without soft agar (Billing, 1969). These methods can be examined and optimised to develop an overall system of identification and extraction of *B. pseudomallei* bacteriophage, for use in downstream applications. These applications include genetic analysis and infection of isolates of *B. pseudomallei* for virulence studies. The described downstream applications require either sufficient quantities of viable bacteriophage, or purified, unsheared DNA. This chapter details the optimisation of methods to produce these objectives.

4.2 Materials and Methods

4.2.1 Identification of bacteriophage-bearing bacteria

4.2.1.1 Isolate selection by cross-spotting

Fifty *B. pseudomallei* isolates were selected from the JCU collection, including soil and clinical samples sourced from Australia and Papua New Guinea. Bacteria were cultured on an Ashdown agar plate for 48 hours (Appendix 1) and subcultured into 10ml LB medium (Appendix 1), which was incubated at 37°C in an orbital shaker (100rpm; Ratek Orbital Mixer Incubator) overnight. A one ml aliquot was removed

and centrifuged to pellet the bacteria (20 000*g* for five minutes; Hettich Universal 16R centrifuge, fixed angle rotor 1614; HD Scientific). The supernatants were collected and used to spot onto plates of isolates as described. One ml of broth of each isolate was spread plated on two LB agar plates (Appendix 1) with excess broth being removed and the plate dried at room temperature for 30 minutes.

Before each use, the applicator was dipped in 70% ethanol and flamed to sterilise it. Supernatants were spotted on each plate using a handmade applicator which was dipped in a microfuge tube holding supernatant, then touched to the plate and held there for five seconds. Supernatant which had attached to the surface of the applicator ran onto the plate and produced a spot. Twenty-six supernatants could be spotted on each plate by this method, requiring only two plates of each isolate to test all other 49 isolates. Plates were allowed to dry and placed in a 37°C incubator overnight. Spot-plate interactions were examined and scored by the following method; 0 = no clearing, 1 = thinning in region of spot, 2 = complete clearing in region of spot, 3 = possible plaque, 4 = clearly defined plaques.

4.2.1.2 Analysis of isolates

Bacterial isolates were screened for the presence of bacteriophage using a mitomycin C assay (Oakey and Owens, 2000). In this assay bacteria were cultured on an Ashdown agar plate for 48 hours and subcultured into 100ml LB medium. This was incubated at 37°C in an orbital shaker (100rpm) until an $O.D._{600nm}$ of approximately 0.2 was reached (Pharmacia LBK Ultrospec3 UV/VIS spectrophotometer). This broth was then aliquoted into 10ml solutions (three replicates each for control and induced sample were included). Induced samples had mitomycin C (Sigma-Aldrich, Castle Hill, NSW, Australia) added to a final concentration of 100ng/ml. When this concentration proved excessive or insufficient, the dose was altered as appropriate to between 20ng/ml and 500ng/ml. The broths were then placed back in the orbital shaker and $O.D._{600nm}$ readings were taken at two hour intervals for a period of 14 hours then eight hourly for a total period of 30 hours and the optical data was plotted (Figure 4.1) . Cell lysis, as indicated by a relative decrease in $O.D._{600nm}$ of the induced sample as compared with the untreated control, was presumptive prophage.

Bacteria found to be resistant to mitomycin C as well as bacteria found to contain inducable prophage by mitomycin C treatment were cultured on Ashdown agar plates and subcultured into LB broth as above. These were incubated at 37°C in an orbital shaker (100rpm) until an $O.D._{600nm}$ of about 0.2 was reached. This broth was then aliquoted into triplicate control and induced samples of ten ml each. Induced samples were placed in Petri dishes and exposed to UV-C (Biorad GS Gene linker) for 1, 2.5, 5, 7.5 and 10 minutes. One minute of UV exposure equated to an exposure of 281mJ. After this time, samples were transferred back into universal bottles and reincubated with protection from light. $O.D._{600nm}$ readings were taken hourly for seven hours then three hourly until 17 hours and then again at about 24 hours. The optical data was then plotted (Figure 4.2). Cell lysis, as indicated by a relative decrease in $O.D._{600nm}$ of the induced samples as compared with the untreated controls and was presumptive prophage.

Confirmation of presumptive bacteriophage was carried out after extraction and concentration of bacteriophage. Three methods were used for this confirmation: Observation with T.E.M. (Section 4.2.7). Observation of plaque forming units present after aliquoting extracted bacteriophage onto a propagating strain (Section 4.2.2.1). Lastly, observation of a bacteriophage DNA restriction enzyme digestion pattern using gel electrophoresis (Section 4.2.3.1).

4.2.2 Extraction of bacteriophage

4.2.2.1 Induction of bacteriophage production

Bacterial broths were grown to an $O.D._{600nm}$ of approximately 0.2 using conditions as for the mitomycin C assay (Section 4.2.1). These samples had mitomycin C added to a final concentration of between 20ng/ml and 100ng/ml depending on sensitivity of individual isolates. These were further incubated until the maximum reduction in O.D. occurred as previously determined or for 18 hours. At this time the broths were centrifuged for ten minutes at 6250*g* at 4°C (Hettich Universal 16R centrifuge, fixed angle rotor 1620A; HD Scientific). The supernatant was decanted and filtered

through 0.45µm syringe filters (Sarstedt; Germany). This bacterially sterile solution (phage filtrate) was then stored at 4°C.

4.2.2.2 Concentration techniques

Five methods were tested for concentration of phage filtrate (Section 4.2.2.1). In each method, yield was determined by the quantity and quality of DNA which could be extracted from the product .

The first method involved the use of a commercial kit for the extraction of DNA from Lambda phage (QIAGEN Lambda Mini kit; QIAGEN). This kit was adapted for use with phage filtrate by starting the protocol at step two of the kit (Appendix 3). The kit was based on polyethylene glycol (PEG) precipitation, sodium dodecyl sulphate (SDS) lysis and low salt binding of DNA to a resin for washing prior to high salt elution. This protocol included DNA extraction. In the remaining four methods, DNA extraction was carried out on the product (filtrate concentrate) using a phenol chloroform extraction method described in Section 4.2.3.

The second method was a standard PEG precipitation method described by Yamamoto *et al.* (1970) where 10% (w/v) PEG 6000 was added to phage filtrate and gently mixed to dissolve. The solution was then chilled at 4°C for at least one hour and the precipitated particles pelleted by centrifuging at 12000*g* for 20 minutes at 4°C (Suprafuge 22; Heraeus Sepatech). Pellets were resuspended in a minimal volume of SM buffer, using a gentle washing technique with a transfer pipette (Sarstedt; Germany). This final volume was labelled filtrate PEG concentrate and stored at 4°C prior to quantitation.

The third method was an adaption of a zinc chloride precipitation method described by Su *et al.*, (1998) in which bacteriophage was precipitated from phage filtrate by the addition of two molar zinc chloride in a volume of one part zinc chloride to 50 parts phage filtrate. This was incubated at 37°C for ten minutes and pelleted by spinning at 4000*g* for five minutes at 20°C. Pellets were resuspended in a minimal

volume of SM buffer, using a gentle washing technique with a transfer pipette (Sarstedt), labelled filtrate zinc concentrate and stored at 4°C prior to quantitation.

The fourth method was an adaptation of a magnesium hydroxide precipitation method as originally used by Vilagines *et al*., (1982) as a second concentration step in collection of bacteriophage from large volumes of water. This protocol concentrates bacteriophage by absorbing negatively charged virions onto a positively charged floc of magnesium hydroxide at high pH (pH 11.5), concentrating the magnesium and solubilising it by neutralisation of pH. A volume of 50mM magnesium hydroxide (typically 100ml) was precipitated by adjusting the pH to 11.5 with sodium hydroxide and stirring for ten minutes. The magnesium hydroxide was pelleted by centrifugation at 2000*g* for five minutes at 20°C and the supernatant discarded. An equal volume of phage filtrate was adjusted to pH 11.5 with sodium hydroxide and added to the magnesium hydroxide pellet which was then homogenised and immediately pelleted again (2000*g*, five minutes, 20°C). The pellet was collected, resuspended in SM buffer to ten percent of the original volume (typically ten ml) and the pH was adjusted to neutral with hydrochloric acid, taking care not to drop the pH below seven. The solution was labelled filtrate magnesium concentrate and stored at 4°C prior to quantitation

The fifth method involved ultracentrifugation of phage filtrate to pellet the bacteriophage (Oakey and Owens, 2000). A 100ml volume of phage filtrate was ultracentrifuged at 200 000*g* for four hours (Optima L-90K ultracentrifuge, TY 50.2 Ti rotor; Beckman Coulter). The supernatant was discarded and the pellet was resuspended in a minimal volume of SM buffer, using a gentle washing technique with a transfer pipette (Sarstedt; Germany). The product (filtrate ultraconcentrate) was stored at 4°C prior to quantitation and further processing.

4.2.2.3 Confirmation by plaque formation

A range of *B. pseudomallei* isolates were selected from the JCU collection based on data collected by cross spotting (4.2.1.1) and grown in LB media overnight at 37°C with shaking at 100rpm. One ml of each isolate was spread plated onto an LB plate

and excess media removed. The plate was allowed to dry for 30 minutes and a six µl drop of concentrated extract (as described in Section 4.2.2.2) was spotted onto the agar. This drop was allowed to air dry and the plate was incubated overnight at 37°C, after which the plate was examined for the production of plaques. Plaque formation in the area of the drop, with no plaque formation on the rest of the plate, was considered confirmatory for the presence of bacteriophage. If complete clearing occurred, the concentrated extract was serially diluted and tested again to eliminate the possibility of the presence of bacteriocins in the extract causing clearance.

4.2.3 Bacteriophage DNA extraction

The efficacy of two methods for the extraction of DNA from concentrated bacteriophage samples were compared. Firstly, a commercial kit for the extraction of DNA from lambda phage (Nucleobond AX; Machery Nagel, Germany) was adapted for use with concentrated bacteriophage samples by eliminating step one of the kit (Appendix 3). The kit was based on PEG precipitation, sodium dodecyl sulphate (SDS) lysis and low salt binding of DNA to a resin for washing prior to high salt elution. This was followed by either isopropanol precipitation (kit method) or ethanol precipitation (Section 3.2.2) and dissolution into 100µl sterile deionised water (ddH_2O). This product was stored at -20°C.

The second method was a phenol chloroform extraction modified for bacteriophage DNA as described by Oakey and Owens, (2000). Modifications to increase yield included increasing proteinase K concentration and increasing digestion time. Prior to lysis, filtrate concentrates (PEG, zinc, magnesium and ultraconcentrate), with and without deoxyribonuclease 1 (DNase 1; Sigma) added to a final concentration of one µg/ml, were incubated for 30 minutes at 37°C to remove any bacterial genomic DNA. The DNase 1 was denatured at 75°C for ten minutes. The filtrate concentrate was then lysed with addition of EDTA, pH 8 (final concentration 20mM), proteinase K (final concentration 300µg/ml) and SDS (final concentration 0.5%) and incubated at 56°C for 1.5 hours in a waterbath. Ribonuclease A was added (final concentration 10µg/ml) and the solution was incubated at room temperature for one hour, then warmed to 56°C for 15 minutes. The nucleic acid was purified using a standard

phenol, phenol/chloroform and a chloroform extraction (Section 3.2.1). The phenol was warmed to 56°C before addition to the lysate, mixed by inversion, and incubated at 56°C for 30 minutes prior to centrifugation. Centrifugation was carried out at 5000*g* for ten minutes for each step. Further concentration was carried out using an ethanol precipitation (Section 3.2.2). In both methods, nucleic acid yield was determined by $O.D._{260/280nm}$ readings and quality of the DNA was observed by using agarose gel electrophoresis (0.8% gel, 80V, 80 minutes).

4.2.3.1 Confirmation by restriction digestion

DNA extracted from bacteriophage, (4.2.3), was digested using a standard restriction digest assay (Sambrook and Russell, 2001) with reference to the usage information supplied with the relevant restriction enzymes. Six enzymes were trialed and selection of restriction enzyme was made on the observation of clear multiple bands on agarose gels. One µg of bacteriophage DNA, as quantified by spectroscopy at 260nm, was used in each digestion. Digestion was carried out for 1.5 hours at 37°C. The digestion pattern was observed on agarose gel (Genesnap; Syngene) stained with ethidium bromide (0.5µg/ml) (Figure 4.3).

4.2.3.2 Comparison of concentration techniques

To determine which of the five concentration methods produced the cleanest and highest quantity of DNA, a volume of filtrate concentrate (five ml) was concentrated and the DNA extracted into a final volume of 500µl. The sole exception to this was the QIAGEN method where the results were calculated by a factor to standardise the initial and final volumes used in comparison with other methods. In all other cases, DNA was extracted from filtrate concentrate using the phenol chloroform method and all DNA products were analysed by $O.D._{260/280nm}$ for yield (ng/µl) and by gel electrophoresis for purity.

4.2.4 Propagation of bacteriophage

Individual *B. pseudomallei* isolates which tested negative by using the mitomycin C method (Section 4.2.1) were screened as a propagating strain for plaque formation with the selected bacteriophage extracts. Any isolate which permitted plaque

formation was further tested by attempting to extract bacteriophage from the isolate using the method described previously (Section 4.2.2). These concentrated extracts were further screened by plaque formation and DNA extraction and digestion protocols (Section 4.2.3). Isolates from which the bacteriophage extracts did not produce plaques on any plate and which, when digested, did not produce DNA banding on agarose were considered not to be carrying viable prophage. A propagating strain (#4) was selected from this group.

4.2.5 Quantitation of bacteriophage

A one ml volume of an overnight broth of the propagating strain was plated onto an LB plate, excess broth was removed and the plate allowed to air dry for a half hour. The bacteriophage solution was titred by serial dilution in SM buffer (Appendix 1) and six µl aliquots were spotted onto the LB plate. Dilutions in initial cases were; neat, 1:2, 1:4, 1:8 and 1:10. After amplification of phage, dilutions were 10- fold to a dilution of 10^{-9}. The six µl aliquots were allowed to air dry and the plate incubated overnight at 37°C. Plaques were counted on plates using the lowest dilution which permitted reliable determination of plaque number and the plaque forming units/ml (pfu/ml) calculated using the following equation;

$$\frac{\text{plaques} \times 1000 \times \text{dilution}}{6} = \text{pfu/ml}$$

Where a 1:10 000 dilution would have a dilution value of 10 000

4.2.6 Amplification and storage of bacteriophage

4.2.6.1 Plate amplification

Initial amplification of bacteriophage was carried out by spotting serial dilutions of the concentrated extract onto a lawn of the selected propagating strain and incubating overnight at 37°C. Growth from plates showing complete clearing was scraped with a plastic loop (Sarstedt: Germany), washed with 500µl SM buffer and pipetted off the plate. This solution was spun briefly to remove cell debris and the supernatant collected. Serial dilutions of 10^{-1} to 10^{-9} were carried out and six µl aliquots were spotted onto a propagation plate. This protocol was repeated until the yield of phage

no longer increased. The resultant amplified bacteriophage was tested by DNA extraction and digestion to confirm purity of the product.

4.2.6.2 Broth amplification

The propagation strain was grown at 37°C in ten ml LB broth to an $O.D._{600nm}$ of 0.1. One hundred microlitres of bacteriophage which had been plate amplified to a concentration of at least $1x10^6$pfu/ml (Section 4.2.6.1) was added and the level of bacteriophage in the supernatant monitored by repeated removal of 500µl aliquots which were spun briefly to remove cell debris and titred (Section 4.2.5). Optimal harvest time was determined by yield of bacteriophage.

Literature suggests that while maximum yield of bacteriophage from lytic phage infection is gained from a 1:1 to 10:1 infection rate (pfu:cfu) (Fischetti *et al.*, 1971; Raina, 1981), in lysogenic systems an infection rate of less than 1:1 is required to maximise the lytic rather than lysogenic lifecycle (Jacob and Wollman, 1959). To test this, a range of inocula were trialed and bacteriophage yield observed.

Harvested bacteriophage was further added to a fresh broth of propagating strain, and the protocol was repeated until bacteriophage levels reached a level suitable for further work, ideally causing complete clearing of the broth. Five ml of the resultant amplified bacteriophage was tested by DNA extraction (Section 4.2.3) and digestion (Section 4.2.3.1) to confirm purity of the product. The rest of the product was stored at 4°C until required.

4.2.7 Characterisation of bacteriophage

Classification to the family level was based on morphological characteristics identified from transmission electron microscopy. Virus size was also determined through T.E.M. A 30µl sample of filtrate ultraconcentrate (Section 4.2.2.2) was transported on ice to the University of Queensland, Department of Microscopy and Microanalysis for examination and analysed using a 1% ammonium molybdate stain. Images of any bacteriophage present were returned and analysed (see Figure 4.8).

4.2.7.1 Determination of viral genome size

Genome size was determined by pulse field gel electrophoresis on a Biorad Chef Mapper XA System using a 1% Tris borate (TBE) gel (M.P. agarose; Roche) in 0.5% TBE buffer. The Chef mapper was programmed using the autoalgorithm function, with a linear program running a range of between 10kb and 150kb (14°C, gradient 6V/cm, included angle 120°, initial.sw.time 47s, final sw. time 12.91s, run time 20 hours). Whole DNA was added to the gel in 1% agarose plugs. Thirty µl of DNA sample consisting of approximately one µg of DNA was added to 30µl of two percent agarose and allowed to cool in plug moulds (Biorad; Australia). Gels were stained for observation under ultraviolet light by insertion into a bath of TBE-ethidium bromide (600µg/ml) solution for 15 minutes. Destaining was carried out by rinsing in 0.5% TBE buffer and soaking in fresh 0.5% TBE buffer for one hour, followed by soaking in 0.1mM magnesium sulphate until the orange colour was removed from the gel (up to one hour). The molecular size of the sample DNA was calculated by comparision to a molecular weight marker (Chef DNA Size Standard Lambda Ladder, Biorad; Australia), using molecular sizing software (Genetools; Syngene).

4.3 Results

4.3.1 Identification of bacteria infected with bacteriophage

4.3.1.1 Cross spotting protocol

Fifty isolates were analysed using this protocol (4.2.1.1), raw data was tabulated (Appendix 2; Collated plaque formation data) and interactions examined. Negative controls consisting of only LB media were determined to have no interaction or some thinning (0 or 1) in each case and all data resulting in these scores was considered to be a media effect and not the result of bacteriophage or bacteriocin activity. Complete clearing (2) was considered to be caused by either bacteriophage or bacteriocin activity, however no halos indicative of bacteriocin activity were observed in these cases. Isolates were scored as possible plaque when pinprick sized indentations in the lawn were observed but could not be clearly identified as plaques. Plaque, possible plaque and complete clearing (4, 3 and 2) were initially grouped as positive for bacteriophage. The qualitative judgements required for this assay meant that data can only be used to develop general impressions of the presence of plaques

and capacity as hosts. Several bacterial isolates also spontaneously produced plaques in the lawn. These included NCTC13178 and #73.

By this method, all isolates tested appeared to show the presence of bacteriophage and most isolates were susceptible hosts to a variable range of bacteriophage extracts. Only two isolates' supernatants produced possible plaques on over 50% of hosts. The supernatant that was indicative for the presence of plaques (score 2,3 or 4) on the greatest number of isolates was #33 with 34 plates out of 50 showing evidence of bacteriophage. Supernatant #30 only showed evidence of bacteriophage activity on one plate (#26), scoring lowest of all isolates.

The isolate which was the best host for production of plaques was NAFC with a score of 37 out of 50 spots producing plaques when spotted on NAFC. The isolate which was the worst host for production of plaques was #82 with no spots producing plaques when spotted on #82.

Thirty isolates from the initial screening of 50 isolates and one clinical isolate (C4) which had not been available on the day of screening, but was from the same region as four environmental samples (E1, E2, E3, E4), were selected for further study (Table 4.1). To ensure a wide range of isolates were studied, selection criteria included; isolates which were related to each other in some way, isolates which were unlikely to produce bacteriophage (scoring 2 or 4<10% of spotting events), isolates which were high producers of bacteriophage (scoring 2 $\geq$ 20% of spotting events), isolates with a broad host range (scoring 2 or 4 $\geq$ 20% of spotting events), isolates with a relatively narrow host range (scoring 2 or 4 between 10% and 20% of spotting events) isolates which were broad hosts for bacteriophage (scoring 2, 3, or 4 $\geq$ 20% of occasions by plate) and isolates which were insensitive to bacteriophage (scoring 2, 3, or 4 <10% of occasions by plate).

While the isolates selected are not an exhaustive grouping of those cross spotted, they do include isolates from each category. Some isolates fit into more than one category, but are only included in one for the purpose of selection.

Table 4.1 Criteria for isolate selection and isolates selected from data collated in Appendix 2

Criteria for isolates	Isolates selected
isolates which are related	NCTC 13178 and NAFC*, E1, E2, E3, E4 and C4**
isolates unlikely to produce bacteriophage	#11 #30 #83 *B. thailandensis*
isolates which are high producers of bacteriophage	NCTC 13179 #69 #73
isolates which have a broad host range	#3 #14 ATCC 23343
isolates with a narrow host range	#4 #6 #9 #12
isolates which were broad hosts for bacteriophage	#2 #5 #7 #8 #10 #19 #29
insensitive to bacteriophage	#13 #18 #82

*NAFC is the isogenic mutant of NCTC 13178 created by G. Ulett (Ulett *et al.*, 2001).
**E1-E4 and C4 are environmental and clinical isolates respectively collected from a single geographic location in Papua New Guinea.

4.3.1.2 Mitomycin C assay

Thirty isolates of *B. pseudomallei* and one isolate of *B. thailandensis* were analysed by the mitomycin C assay, (Table 4.2). An example of a presumptively positive mitomycin C assay with isolate #7 (Figure 4.1a) and an example of a presumptively negative mitomycin C assay with isolate C4 (Figure 4.1b) are demonstrated. Almost half of the isolates tested (41.9%) were presumptively positive by this assay. Some isolates, such as *B. thailandensis*, required a lower dosage of mitomycin C (20ng/ml) as they were found to be very sensitive to the mitogen. In contrast, other isolates did not appear to be sensitive to mitomycin C at all (#14 and C4) as growth was not retarded irrespective of the dose added (up to 500ng/ml).

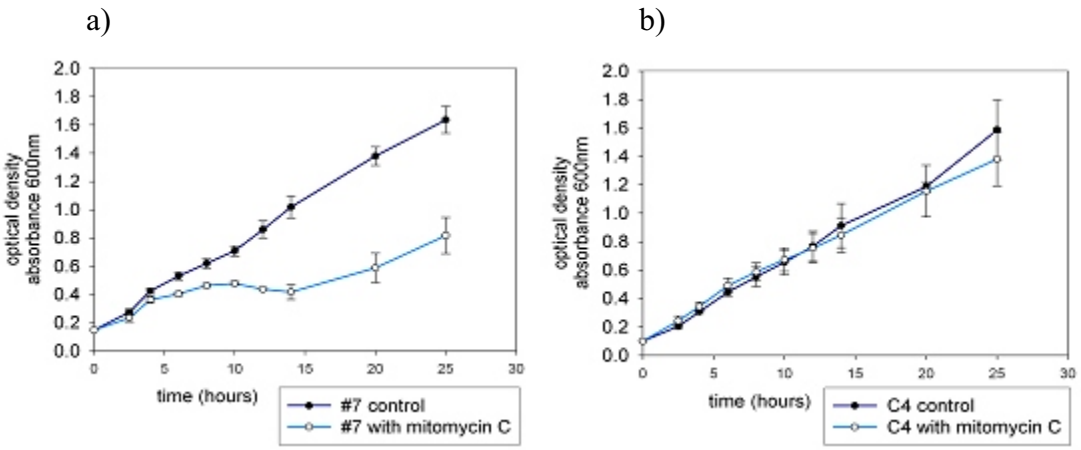


Figure 4.1 Growth curves of a) isolate #7 and b) isolate C4, with and without the addition of mitomycin C; standard error bars included (three replicate assays).

4.3.1.3 Ultraviolet assay

Isolate C4, which was previously found to be insensitive to the mitomycin C assay was tested for the presence of bacteriophage using UV light. A range of exposures was trialed and while all exposures slowed the growth of the organism, none were found to cause a reduction in $O.D._{600nm}$. A reduction would be expected in the presence of bacteriophage undergoing the lytic cycle (Figure 4.2b) . This lack may have been due to the absence of any bacteriophage, however the confirmatory plaque formation assay did indicate the presence of a bacteriophage. To clarify these conflicting results isolate NCTC 13178, which was found by mitomycin C assay and confirmatory assays to carry lysogenic bacteriophage, was tested using the ultraviolet light method (Figure 4.2a). This isolate was a highly virulent clinical strain of *B. pseudomallei*, collected for the JCU culture collection, characterised and then sent to the National Collection of Type Cultures (Britain) to have its phenotypical aspects standardised, after which it was given the code NCTC 13178. This isolate also showed no evidence of the presence of bacteriophage using the UV method. The exposed samples did not grow at the same rate as the unexposed samples, but this is not an actual reduction in optical density. It can be explained by stress to the bacterial cells causing reduction in growth rather than lysis by bacteriophage. Several other isolates were also tested (data not shown) and in each case no evidence of bacteriophage was found. It was surmised that either the equipment for UV exposure

was inadequate to the task, or *B. pseudomallei* strains tested were not sufficiently sensitive to this type of stressor. As the UV exposure did cause stress to the cells, resulting in lower growth, insufficient UV irradiation due to absorption of UV by the broth is unlikely. However, use of a positive control such as a lambda lysogen known to be sensitive to UV irradiation may have been helpful in clarifying this. UV irradiation of broth has also been used successfully for induction of bacteriophage from *B. thailandensis* previously (Woods *et al.*, 2002) albeit with a three ml rather than ten ml broth sample. Given the insufficient or non-existent phage yield, the UV method was not used further.

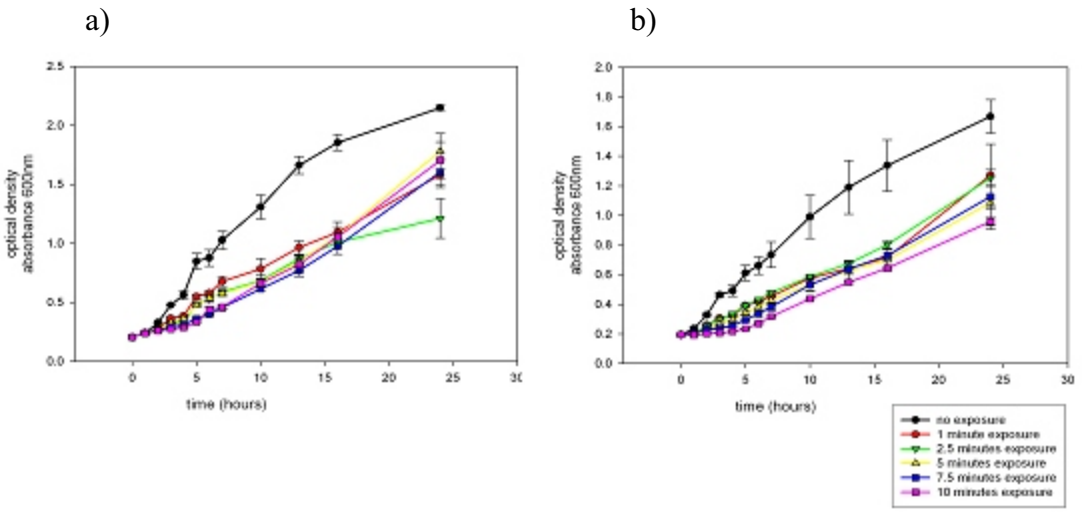


Figure 4.2 Growth curves of *B. pseudomallei* a) isolate NCTC 13178 and b) isolate C4, given various exposures to ultraviolet light (254nm); standard error bars included (three replicate assays).

Table 4.2 Analysis of *Burkholderia* isolates for the presence of bacteriophage; presumptive and confirmatory data. Where a result was borderline or indistinct, data is indicated by '+?'.

Bacterial Isolate	T.E.M. analysis	Mitomycin C presumptive positive	Plaque formation assay	DNA digest assay	Gold standard (MitomycinC, Plaque and DNA digestion assays positive)
NCTC 13178	+	+	+	+	+
NAFC	+	+	+	+	+
2	+	+	+	-	-
3		-	-	+	-
4		-	-	-	-
5		-	+?	-	-
6		-	+	-	-
7	+?	+	+	+	+
8		-	-	-	-
9		-	+?	-	-
10		-	+	-	-
11		-	-	-	-
12		-	-	-	-
13		-	-	-	-
14		-	+	+	-
18		-	+	+	-
19		-	+	-	-
29		+	+	+	+
30		-	+	+	-
73		+	+	+	+
82		+?	+	+	+
NCTC 13179		+	+	+	+
C4	+	-	+	-	-
ATCC 23343		+	-	+	-
#69		-	-	-	-
#83		-	+	-	-
E1		+	+	+	+
E2		+	+	+	+
E3		+	+	+	+
E4		-	+	-	-
B. thailandensis	+	+	+	+	+

Table 4.3 Average dimensions of bacteriophage observed by T.E.M.

Bacteriophage	width of head (nm)	standard deviation	length of tail (nm)	standard deviation	number of images used for calculation
NCTC 13178	63	4	115	11	21
NAFC	57	7	108	15	12
#2	58	3	123	3	3
#7	58	1	no tails seen		4
C4	48	3	163	12	7
B. thailandensis	57	6	113	15	20

4.3.1.4 T.E.M. analysis

Six isolates were sent to the University of Queensland for T.E.M. imaging. Multiple images were used to determine bacteriophage dimensions and standard deviations calculated based on this (Table 4.3). NCTC 13178 could not be differentiated from NAFC, #2 or *B. thailandensis* based on these measurements. The images from #7 only contained empty heads, so this is considered a borderline positive in Table 4.2. Only C4, which was isolated in Papua New Guinea, shows distinct morphology. All samples other than #7 could be clearly identified as *Myoviridae* (See Appendix 6).

4.3.1.5 Confirmation of bacteriophage by restriction digest

Thirty-one isolates were tested by restriction digest and only one presumptively positive isolate (#2) was determined to be negative using this assay. However, there were several weak positives due to low DNA yields. Four presumptively negative isolates (#3, #14, #18, #30), as determined by the mitomycin C assay also showed a positive result using this method.

Six restriction enzymes were initially trialed to select an appropriate restriction enzyme for bacteriophage from *B. pseudomallei*. These were selected as they were known to cut lambda phage into between zero and 30 bands. The restriction enzyme *Eco*RV was found to produce the clearest pattern of bands when tested with isolate NCTC 13178 and used in all subsequent DNA confirmatory tests. While few morphological differences were noted by T.E.M. (Appendix 6), the restriction

patterns observed were found to be different in most digested isolates. Exceptions to this were E1 which was identical to E2, NCTC 13179 which was identical to #82 and NCTC13178 which was identical to NAFC. Isolates with identical restriction patterns did not have identical plaque production patterns (Appendix 2). This indicates that in most cases, isolates of *B. pseudomallei* are carrying a population of heterogeneous bacteriophage.

At least one µg of DNA was required to guarantee visualisation of patterns of bands from restriction digests. Any isolate producing low yields of bacteriophage would also produce low yields of DNA and potential false negatives. For this reason, restriction digestions were only carried out on concentrated extracts (filtrate ultraconcentrates). The bacterial genome of NCTC 13178 showed no signs of digestion (Figure 4.3), while the concentrated bacteriophage DNA did. This confirmed the protocols for separating bacteriophage from genomic DNA were effective and that the digested DNA was from bacteriophage.

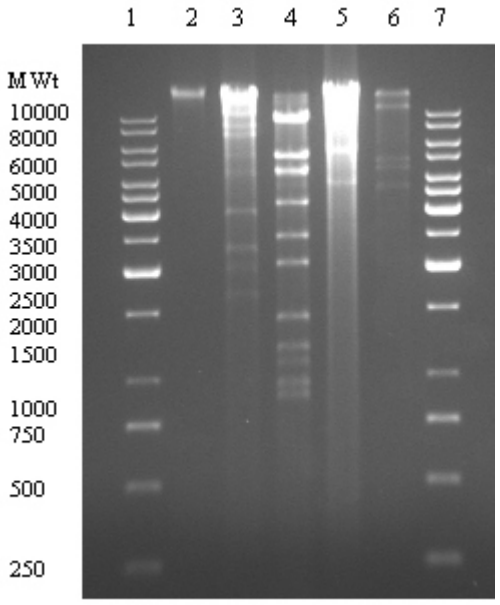


Figure 4.3 *Eco*RV digest of 1µg DNA extracts from NCTC 13178 bacterial genome and bacteriophage isolated from various *Burkholderia* bacteria. Lane 1, 7; Progen Generuler 1kb DNA ladder, lane 2; NCTC 13178 bacterial DNA, lane 3; NCTC 13178 bacteriophage (BupsΦ1) DNA, lane 4; #29 bacteriophage DNA, lane 5; #82 bacteriophage DNA, lane 6; *B. thailandensis* bacteriophage DNA

4.3.1.6 Confirmation by plaque formation

Of 31 isolates tested, only one (ATCC 23343) which was presumptively positive for bacteriophage (mitomycin C assay) failed to produce plaques. Eleven presumptively negative isolates (#5, #6, #9, #10, #14, #18, #19, #30, C4, #83, E4) also produced plaques.

To increase the chances of using a compatible host, multiple hosts were used in each case (Appendix 2). Of the samples tested, only one isolate (ATCC 23343), which was confirmed positive for the presence of bacteriophage by DNA digest, would not form plaques on any of the hosts utilised.

Plaques were often observed to have a very small diameter and were always less than one millimetre in diameter. It should be noted that isolates NCTC 13179 and #82 produced different patterns of plaques, as did E1 and E2 (Appendix 2), indicating that, according to this protocol, these were heterogeneous bacteriophage.

4.3.1.7 Statistical analysis of assays

While T.E.M. was the gold standard of choice, the cost and accessibility of this assay were prohibitive. Instead, a result was considered to be an unconditional positive (gold standard) when all three regularly used assays (mitomycin C, plaque and restriction digest) gave a positive result. Based on this new gold standard, the data was analysed to determine sensitivity, specificity, positive predictive value (P.P.V.), negative predictive value (N.P.V.) and accuracy of each assay and combination of assays (Table 4.4). When a combination of assays was analysed, the assays were considered to be positive when every assay was positive and negative when one or more assays were negative. Calculations and equations are in Appendix 5. Sensitivity and N.P.V. are artificially 100% due to the nature of the gold standard. As such, comparison can only be between specificity, accuracy and P.P.V. In isolation, the mitomycin C assay was the most specific, accurate and gave the highest P.P.V. Over all, any two assays of which the mitomycin C assay was one gave the best results.

Table 4.4 Statistical analysis of result of the results of Table 4.2 based on a gold standard of three positive assays being taken as positive for bacteriophage. Data is recorded as a % . P.P.V. is positive predictive value, N.P.V. is negative predictive value.

	sensitivity	specificity	P.P.V.	N.P.V.	accuracy
mitomycin C assay	100	90	84.62	100	93.55
plaque assay	100	40	47.83	100	61.29
restriction digest assay	100	75	68.75	100	83.87
mitomycin C and plaque assay	100	95	91.67	100	96.77
mitomycin C and restriction digest assay	100	95	91.67	100	96.77
plaque and restriction digest assay	100	85	78.57	100	90.32
mitomycin C, plaque and restriction digest assay (gold standard)	100	100	100	100	100

4.3.2 Extraction of bacteriophage

4.3.2.1 Induction of bacteriophage production

B. pseudomallei isolate NCTC 13178 was used for extraction trials. Phage filtrate was found (by plaque formation assays) to yield approximately $1x10^3$ pfu/ml. This yield of bacteriophage was insufficient for downstream applications such as DNA extraction and infection and required a further concentration step. From this point on this bacteriophage was allocated the name BupsΦ1.

4.3.2.2 Concentration techniques

The QIAGEN kit was found to produce the least amount of DNA by $O.D._{260nm}$ and purity could not be determined on a gel due to low yield (Figure 4.4). The magnesium hydroxide method produced the second lowest yield and was also not visible on gel electrophoresis. Concentrated bacteriophage produced by the magnesium hydroxide method was also found to be non-infective by plaque spotting (data not shown), further indicating that either structural damage had occurred or the presence of magnesium hydroxide or unfavourable pH levels were affecting the infection process.

In all other cases, the protocols (Section 4.2.2.2) were found to produce a clean band of DNA of greater than ten kb and a smear of sheared DNA less than 250 bases in size. Both the PEG method and the zinc chloride method produced intermediate levels of clean DNA.

The ultracentrifugation method produced the greatest yield of DNA in the ten kb region and the lowest amount of sheared DNA (Figure 4.4). This method had the advantage of not adding any additional reagents which could affect the infectivity of the bacteriophage hence it was selected for subsequent use.

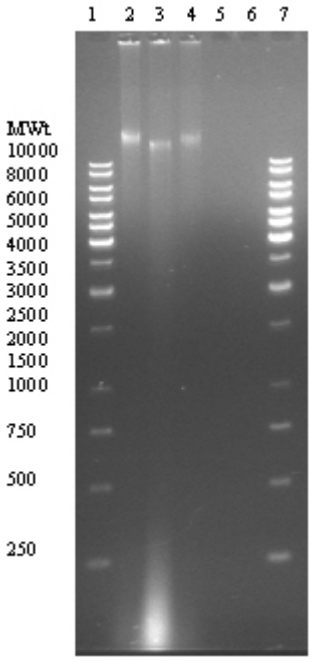


Figure 4.4 Comparison of concentration technique as shown by quantity and quality of DNA extracted from concentrated product. Lane 1, 7; Generuler 1kb DNA ladder (Progen), Lanes 2-6 concentrated DNA extracts: lane 2; ultracentrifuged, lane 3; zinc chloride precipitated, lane 4; PEG precipitated, lane 5; magnesium hydroxide precipitated, lane 6; QIAGEN kit extraction

Addition and neutralisation of DNase 1 prior to lysis of the bacteriophage capsid, for the purposes of removing contaminating genomic DNA, did not remove the smeared band of sheared DNA, but did lower DNA quantity and intensity of the upper band. This loss may have been due to mechanical damage caused by the precipitation methods, permitting entry of DNase 1 into the capsid and destruction of the DNA.

The upper clear band was determined to be bacteriophage DNA by restriction digest. The restriction enzymes used produced no digestion with host genomic DNA and a clear band pattern with bacteriophage DNA.

4.3.3 Extraction of bacteriophage DNA

The Nucleobond AX kit was found to be difficult to operate as alteration of pH of the bacteriophage concentrate to pH 6.3 was required. As the maximum volume of culture with this kit was only ten ml, a standard pH probe was not practical so pH strips (Neutralit pH5-10; Merck) were used to monitor the pH change. This proved expensive as each strip could be used once only, and over-acidification was a problem. To minimise this problem, prior to resuspension, media used to resuspend pellets of bacteriophage were changed to appropriate pH and potassium chloride conditions. However, monitoring and adjustment was still required. As the kit was designed for extraction directly from a high yield broth, working with concentrated samples of very low volume required modification in volumes and weights of the media in the kit. Measurement of extraction matrix was particularly difficult, requiring use of a fine analytical balance (four decimal place basic balance; Sartorius). Yield with this kit was found to be very low and investigation indicated that DNA was lost in the isopropanol precipitation step. Use of fresh isopropanol, or use of an ethanol precipitation step instead, did not increase yield and this method was discarded as being too problematical.

The phenol chloroform method was found to give sufficient yield of DNA for downstream applications such as restriction digestion or cloning. However, the quality of the DNA was poor as a large proportion of the DNA consisted of sheared fragments. These fragments may have included genomic DNA remaining from original lysis of the bacterial cells. Removal of these fragments was carried out by processing the DNA through a PCR clean up kit (Wizard SV gel and PCR clean up system; Promega). The yield of DNA through the kit was typically ten percent of that before processing. This protocol had advantages over the Nucleobond AX kit as adjustements for different volumes of sample were easily made. To minimise use of phenol, large volumes of broth containing bacteriophage (phage filtrate or filtrate

concentrate) could be freeze dried and resuspended in smaller volumes of sterile distilled water without compromising yield or quality. This method was found to be robust and was used for all subsequent DNA extractions.

4.3.4 Quantitation of bacteriophage

Initial quantitation studies using a soft overlay method of plaque formation were ineffective as it was not often possible to observe plaques. Direct plating of plaque extract onto a lawn of host bacteria permitted the observation of plaques. These plaques were less than one millimetre in diameter and slightly turbid.

4.3.5 Propagation of bacteriophage

4.3.5.1 Selection of a propagation host

The data collected from the plaque assay (Appendix 2) indicated that some bacteriophage, including NCTC 13178, *B. thailandensis* and #73, had a wide host range. In addition, some extracts did not indicate the presence of bacteriophage on any host. It cannot be concluded that these isolates are negative for the presence of bacteriophage by this assay alone as it is possible the negative result was due to the lack of a compatible host being available for testing the extracts. For the purpose of this study, confirmation of the absence of bacteriophage requires two or more assays to indicate that there is no bacteriophage present in the extract and there must be are no conflicting positive results. Isolates #4, #8, #11, #12, #13, and #69 were found to be negative by all assays (Table 4.2). Isolate #4 was selected as a potential propagation strain.

4.3.5.2 Selection of bacteriophage for propagation

Initial selection of a bacteriophage for further analysis was based on compatibility with propagation strain #4, and broad infectivity. Eight bacteriophage isolates fitted this selection criteria; bacteriophage extracted from NCTC 13178, #29, #30, #73, *B. thailandensis,* E1, E2 and NAFC. From this selection, the bacteriophage extracted from NCTC 13178 (BupsΦ1) was chosen for further work as the bacterial isolate is highly virulent, is of interest with respect to virulence studies and has been

extensively characterised at James Cook University (Leakey *et al.*, 1998; Ulett *et al.*, 2000a; Ulett *et al.*, 2000b; Ulett *et al.*, 2001).

Bacterial isolate # 4 was infected with concentrated extract BupsΦ1 and the mitomycin C assay produced positive results. These were confirmed by the plaque assay (on propagation strain #4) and restriction digest.

4.3.6 Amplification and storage of bacteriophage

Initial concentrations of bacteriophage were too low to cause any increase in yield when inoculated into a broth system as the infection rate was too low. A plate system, such as plaque formation, has a lower concentration of host cells, resulting in a higher infection rate and easier collection of progeny bacteriophage.

4.3.6.1 Plate amplification

Initial trials of standard soft agar overlay as used in previous work (Hayashi *et al.*, 1990) were not successful in amplifying phage. Plaques were less than one millimetre in diameter when using direct bacterial plating and could not be observed in soft agar. It was assumed that the depth of the soft agar and the presence of bacteria in it masked the plaque, thus confounding their harvesting. For this reason all subsequent work was carried out using the direct agar plating technique.

Serial plate amplification was carried out and the yields monitored for two weeks. The yield increased about one log amplification per cycle. However, loss of yield was observed when cell debris was not immediately removed from scraped plaques. This loss can be explained by binding of the bacteriophage to cell wall debris. The maximum yield obtained was $1x10^{10}$pfu/ml in a volume of 500µl. Further amplification attempts did not increase yield. This protocol, while useful in initial amplification, did not produce a high enough yield for downstream applications and so broth amplification was attempted.

To confirm only a single type of bacteriophage was amplified, single plaques were extracted from a plate and processed for four cycles of extraction and plating of

single plaques. The pure bacteriophage were processed via DNA extraction and digestion using *Eco*RV and the single plaque products were found to produce an identical band pattern to the mass amplification product, confirming the presence of only one type of bacteriophage.

4.3.6.2 Broth amplification

Initial inoculation of bacteriophage from plate amplification into a chilled LB broth of #4 *B. pseudomallei* (at $O.D._{600nm}$ 0.1, equating to $1x10^8$ cfu/ml) was carried out and yield monitored hourly (Figure 4.5). The curve demonstrated that the initial inoculation of bacteriophage ($1.29x10^5$ pfu/ml) dropped as bacteriophage infected the cells, and then rose again as cells begin to lyse. The first visible evidence of lysis in the media (clumping of dead cells in strings) occurred at seven hours.

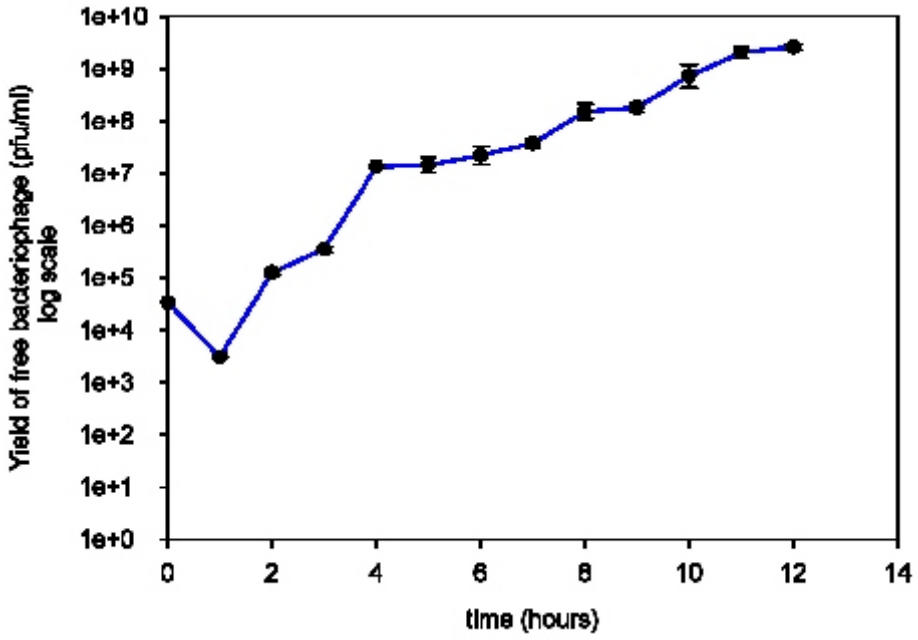
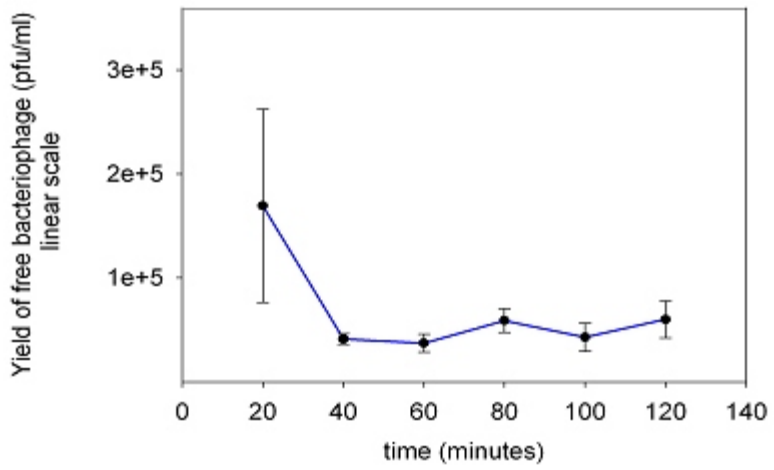


Figure 4.5 Initial broth amplification experiment; yield of free bacteriophage in a broth of *B. pseudomallei* over time. Yield in log scale, standard error bars included (three replicate assays).

When $6.7x10^5$ pfu/ml of bacteriophage was inoculated into prewarmed broth, infection and lysis (as shown by yield of free bacteriophage) can be observed in 80 minutes. Infection and lysis was shown by a drop in free bacteriophage (infection) followed by a rise in free bacteriophage (lysis). On a log scale (Figure 4.6b) this was difficult to observe. The inoculating dose was subtracted from Figure 4.6a to aid visualisation of small fluctuations in phage yield.

a)



b)

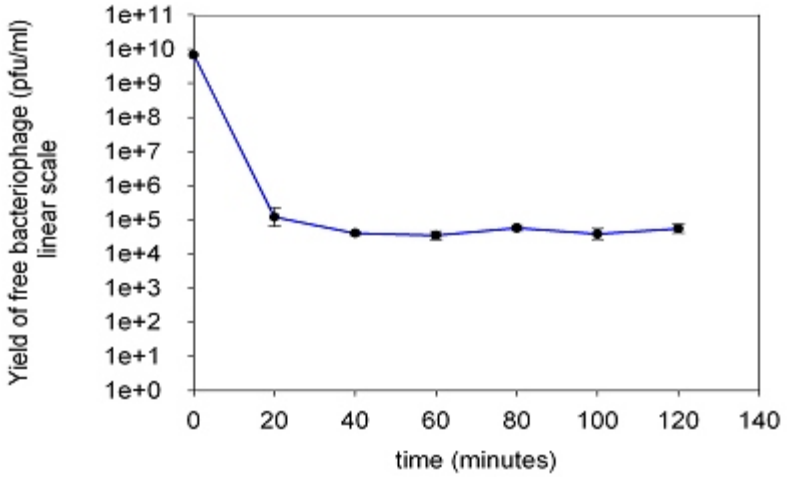


Figure 4.6 Bacteriophage yield up to two hours after inoculation with $6.7x10^5$ pfu/ml bacteriophage into 10ml of $1x10^8$cfu/ml #4 *Burkholderia pseudomallei*. a) linear scale, b) log scale; standard error bars included (three replicate assays).

To test for the optimal infection dose of BupsΦ1 in *B. pseudomallei* #4, a range of inoculum doses were trialed (Figure 4.7). Bacteriophage yield was observed at seven hours, this being the time at which lysis had been seen in initial studies, and 11 hours,

this being the time at which log increases in yield had ceased in initial studies (Figure 4.5). Initially the broth contained $1.6x10^8$cfu/ml (#4) *B. pseudomallei* in ten ml. At seven hours, there was a logarithmic increase in yield as the dose increased logarithmically, up to an infection rate of approximately 1:200 pfu:cfu (inoculating dose $6.7x10^6$ pfu). After this, while yield increased, it did so non-logarithmically. Delaying harvest to eleven hours resulted in increased yields for lower inoculums up to $6.7x10^5$ pfu (1:2000) after which increases in yield were no longer logarithmic and were no greater than those harvested at seven hours.

Lysis, as seen by clearing of the media, was observed to have occurred by seven hours in broths inoculated with $6.7x10^6$ - 10^9 pfu, with most clearing in the 10^7 (1:20) and 10^8 (1:2) inoculums. By eleven hours, clearing in these samples had progressed further, with 10^7 and 10^8 still being the most cleared.

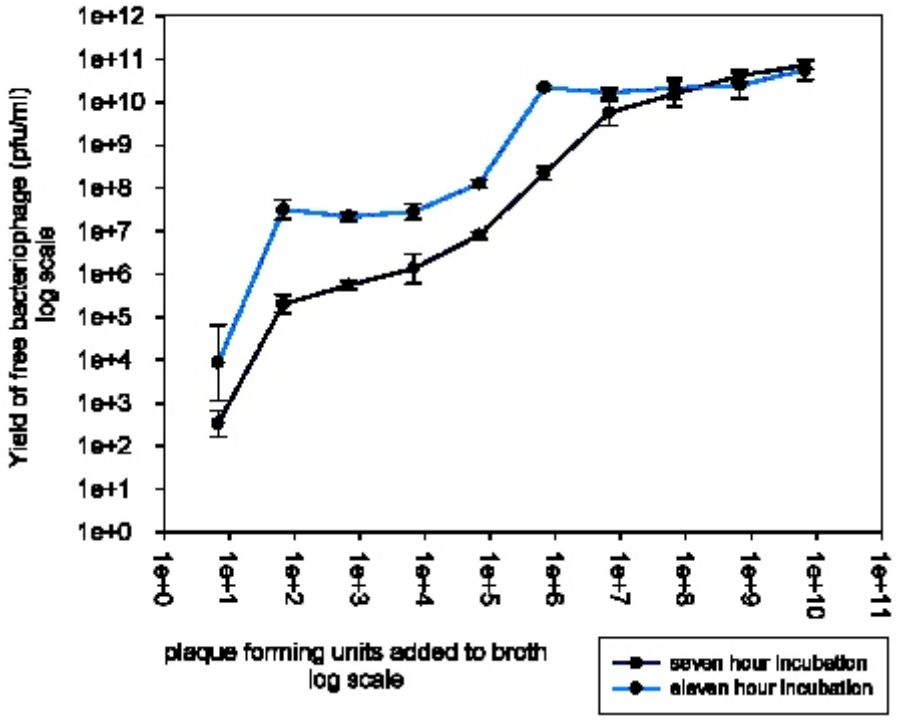


Figure 4.7 Bacteriophage yield at seven and eleven hours post-inoculation of a range of doses of BupsΦ1 into ten ml of #4 *B. pseudomallei* at $1.6x10^8$ cfu/ml; standard error bars included (three replicate assays).

It was determined that a 1:1 dose of pfu:cfu should be used with harvest carried out after lysis was observed (about seven hours) as no increase in yield occurred when the sample was left for a longer period.

Once inoculum size was optimised, a larger volume of broth containing #4 *B. pseudomallei* was inoculated to produce enough bacteriophage for downstream applications. One hundred millilitres of broth at $O.D._{600nm}$ of 0.1 was inoculated with a 1:1 dose of BupsΦ1. Lysis was seen at 7.5 hours, at which stage the broth was filtered and bacteriophage levels titred. The yield, as determined by plaque assay (Section 4.2.5), was $1x10^{11}$pfu/ml.

To ensure purity of product, a sample was tested by DNA restriction digest and compared to an original sample. The samples produced identical bands. Repeated attempts to use higher inoculation ratios in large volumes failed to cause lysis.

4.3.7 Characterisation of bacteriophage

The bacteriophage BupsΦ1 extracted from NCTC 13178 can be can be classified into the family *Myoviridae* based on the presence of an icosohedral head, a neck/collar region and a sheathed rigid tail (Figure 4.8). Based on 21 images, the bacteriophage had 115 ±11nm long tail and a 63±4nm wide icosahedral head.

The bacteriophage nucleic acid was confirmed to be double stranded DNA by restriction digest and digestion with DNase 1. The DNA was determined to be linear by digestion with exonuclease (Figure 4.9). Using Genetools software, on the pulse field gel of the undigested BupsΦ1 genome, the genome was determined to be 55142 bases (or 55.1 kb) long (Figure 4.10).

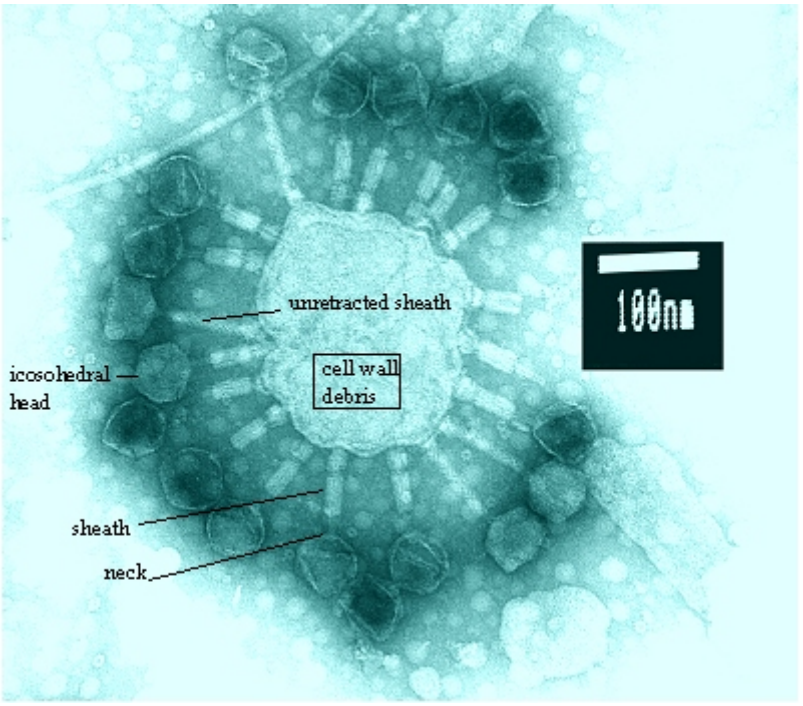


Figure 4.8 Transmission Electron Microscopy view of BupsΦ1

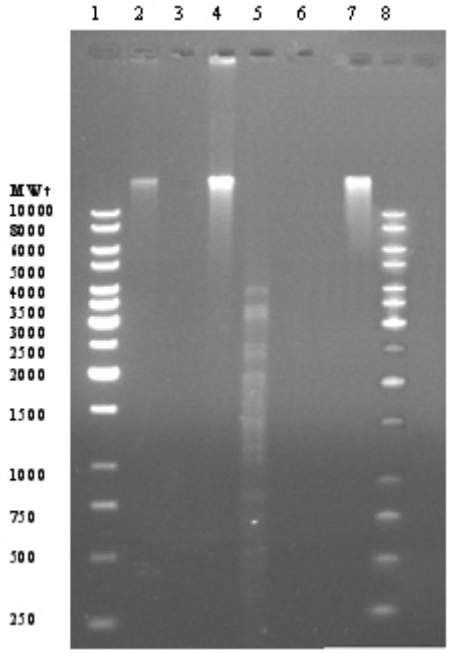


Figure 4.9 Digest of 250ng of BupsΦ1 DNA for confirmation of nucleotide type using the following enzymes; lane 2; no digestion (125ng), lane 3; deoxyribonuclease 1 (Sigma), lane 4; ribonuclease A (Sigma), lane 5; Sal1 (Promega), lane 6; Bal31 exonuclease (Promega), lane 7; S1 nuclease (Promega). This Figure is a composite of two gels with the left marker and lanes 1-6 being one gel and the right marker and lanes 7-8 being another gel. Backgrounds were merged using Corel™ photo-paint 11. The marker used (lanes 1, 8) was Generuler 1kb DNA ladder (Progen).

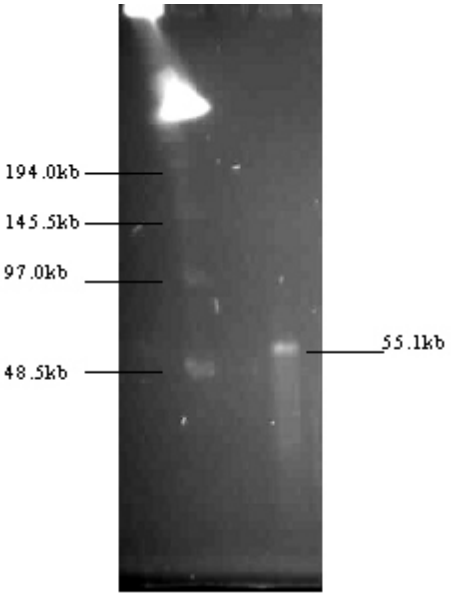


Figure 4.10 Pulse field gel of BupsΦ1 genome; lane 1 Lambda DNA ladder (Biorad, USA), lane 2 BupsΦ1 DNA extract

4.4 Discussion

4.4.1 Preliminary screening

Examination of the cross-spotting table (Appendix 2) showed that no two isolates had the same cross-spotting profile. This was similar to results reported by Gamage (2004). Comparison of grouping of isolates (Table 4.1) with the presence or absence of bacteriophage (Table 4.2) showed some unexpected differences. No broad host range isolates as identified in Table 4.1 met the gold standard, although all isolates had at least one presumptive assay positive. Most other results were expected, with all isolates with narrow host ranges classified negative against the gold standard. This indicated that those hosts which appeared to show the presence of bacteriophage may have been self-lysing, possibly due to stressors applied by addition of LB medium containing toxic byproducts of growth. While these would also be present in the broth of the host, in the position of the spot, extra media is applied and this dose could be enough to trigger lysis. Of those isolates selected as unlikely to produce bacteriophage, all except one were clearly negative for bacteriophage. The positive isolate, *B. thailandensis*, has also been previously determined to be positive for bacteriophage (Woods *et al.*, 2002). The variation in results between Table 4.1 and

Table 4.2 could be due to the possible inconsistences when evaluating the cross spotting assay by eye. As such, this method should not be used on its own for detection of bacteriophage, but is useful as an initial screen to detect isolates of interest to study further.

One set of isolates of interest are NCTC13178 and NAFC. NAFC is an isogenic mutant of NCTC 13178 with a different colonial morphology on Ashdown agar (Ulett *et al.*, 2001). These isolates have similar but not identical spotting patterns, though this variation may be in some part due to observational fluctuations. In addition, these isolates have very different characteristics as hosts. NCTC 13178 was a poor host (scoring 2, 3 or 4 seven times out of fifty) and NAFC was the broadest host examined (scoring 2, 3 or 4 thirty-one times out of fifty). The colonial morphological differences between these isolates indicates that variation in host capacity may be due to changes in cell surface structure, making the NAFC isolate more permissive to bacteriophage.

4.4.2 Methods of detection of lysogens

The first presumptive bacteriophage detection test, mitomycin C, was found for the most part to be an effective trigger for lysis while UV irradiation was not. Literature suggests that the bacteriophage yield using the UV method can be lower than that of the mitomycin C method (Loessner *et al.*, 1991) and that in some cases bacteriophage are not UV inducible at all (Calendar, 1970; Barksdale and Arden, 1974). If yields were too low, observations using optical density may be too insensitive to detect lysis. There is, however, some conflicting work which indicates both methods are approximately equal in efficiency, but not necessarily on the same isolate (Jiang and Paul, 1998). Previous work using strains of *B. thailandensis* (Woods *et al.*, 2002) showed variable success with UV induction, with some strains not being inducible, while others showed a log increase in bacteriophage yield. It should be noted that some prophage cannot be induced by mitomycin C or UV-C (Calendar, 1970). Therefore even the use of both methods cannot, on their own, eliminate the

possibility of the bacteria being a lysogen. This phenomenon was observed in the examination of isolate C4.

As observed in Table 4.2, not all assays produce the same result. This can be explained by the nature of these assays.

The mitomycin C assay is dependant on the triggering of the lytic cycle by the bacteriophage. If the production of bacteriophage is low, the growth of the isolate may outpace its lysis, hence the lysis is not visible as a reduction in optical density. The confirmatory assays, however, are carried out on concentrated broth extract and as such, are more likely to permit the observation of bacteriophage. It is also possible for bacteriophage to be produced at low levels during exponential growth without specific induction (Billing, 1969). In this case, even if induction by mitomycin C did not work (Calendar, 1970) it may be possible to identify these spontaneously produced bacteriophage by one or more of the confirmatory assays.

T.E.M. analysis is considered to be the most sensitive of the assays, as a single bacteriophage can be observed. This assay is also the most expensive of the three assays, with analysis of a single isolate being approximately AU$100. Due to the expense, this assay was not used widely in this study.

On occasion, either the plaque formation assay or DNA digest assay produced negative results when other assays were positive. One explanation for negative results in the plaque formation assay is the use of inappropriate host strains.

Multiple *Burkholderia* isolates have been used previously to screen *Burkholderia* bacteriophage (Leclerc and Sureau, 1956; Denisov and Kapliev, 1991). In all cases, *B. pseudomallei* isolates have been found to be variably sensitive to the extracted bacteriophage as seen in this study. In each case, the concentrated extract was spotted onto multiple strains of *B. pseudomallei*. In most cases, while some hosts produced plaques, others did not (Appendix 2), indicating not all *B. pseudomallei*

strains are appropriate hosts for all bacteriophage extracted from other *B. pseudomallei* strains. The host range of temperate bacteriophage is determined by multiple factors. Bacteriophage require cellular components for replication and so become specialised to a compatible bacterial species, with identification of this species by cell surface receptors. Within a host species, divergent receptors create barriers to interstrain infection (Williams *et al.*, 2002). An example of this is the relationship of the bacteriophage ΦCTX with its host *Pseudomonas aeruginosa* where the bacteriophage requires a particular LPS core structure for infection. *P. aeruginosa* strains have been found to be heterogeneous for this core structure, and hence the bacteriophage will not infect all strains (Yokota *et al.*, 1994).

Another complication of the plaque formation assay is that it requires the bacteriophage to spontaneously enter a lytic cycle in its new host. Given that the bacteriophage had to be triggered into undergoing the lytic part of the lifecycle when originally extracted, it may, on entering most other strains, enter lysogeny. The small, slightly turbid nature of the plaques in this study is typical of lysogenic systems. Plaques produced from lysogenic bacteria are typically "cloudy" as yield of spontaneously lysing bacteria is low (Billing, 1969) and this can make them very difficult to observe, particularly when the plaques are naturally very small. Alternatively, if the potential host is already a lysogen, the prophage present may protect the bacteria from superinfection (Jacob and Wollman, 1959) and preclude entry to the bacteriophage being tested. Any of these situations would result in a false negative result.

Traditionally the host strain used for plaque formation assays with *B. pseudomallei* bacteriophage is *B. mallei* (Denisov and Kapliev, 1991; Denisov and Kapliev, 1995; Woods *et al.*, 2002). The advantage of using *B. mallei* as a host is that it is believed to be universally sensitive to *B. pseudomallei* bacteriophage (Denisov and Kapliev, 1991). This bacterium was not used in this assay due to the highly infective nature of *B. mallei* and the seriousness of the disease caused in humans (Howe and Miller, 1947; Smith, 1957; Srinivasan *et al.*, 2001). The increased difficulty of the methods

used in this study was considered an acceptable trade off for the lower risk of human infection.

In some cases, the plaque formation assay indicated evidence of bacteriophage when DNA extraction did not. Successful DNA digests were dependant on a high yield of bacteriophage. If yields of bacteriophage were low, observation would be difficult and false negatives could be produced. It was assumed that in some cases there was insufficient DNA for visualisation, but sufficient bacteriophage for plaque assay. This was based on the calculation that only 167 pfu/ml (one pfu from six µl of extract spotted on a host plate) are needed for visualisation in the plaque assay. Previous work has shown that the levels of bacteriophage extracted from spontaneous lysis of *B. pseudomallei* can vary dramatically (Denisov and Kapliev, 1991) so it is possible that the same could be true with induced lysis.

No test is perfect, therefore a test or combination of tests is set as the gold standard (knowing that some isolates will be missed) and then accuracy is measured against the gold standard. As a gold standard, any isolate which had three positive assays was considered to be positive for the presence of bacteriophage. For the purposes of identifying a potential bacteriophage-free host, all three assays were required to be negative.

The definition of the gold standard used to analyse these assays means that statistical analysis of sensitivity and negative predictive value is meaningless as the conditions for any result other than 100% cannot be met. However, specificity, P.P.V., and accuracy can be used to select the best assay or group of assays for the purposes of identifying bacteria containing bacteriophage. In each case, the level of specificity, P.P.V. and accuracy were correlated, with low specificity matching low P.P.V. and low accuracy and *vice versa*. The low specificity of the plaque assay is understandable as bacteriocins or antibiotics present in the extract could potentially be mistaken for plaques, or hide the presence of bacteriophage. Plaques were also typically small and cloudy and sometimes difficult to differentiate. The restriction

digests were an improvement with respect to specificity, but still not acceptable in isolation. The mitomycin C assay was highly specific and addition of either plaque or restriction digest assay improved the specificity to 95%, the positive predictive value to 91.67% and the accuracy to 96.77%. Either of these combinations would be acceptable for further work.

4.4.3 Concentration methods

Of the concentration protocols trialed, only ultracentrifugation was found to produce viable bacteriophage from which relatively clean, high molecular weight DNA could be extracted. The bacteriophage was shown to be sensitive to mechanical damage caused in concentration. Structural damage to the bacteriophage caused a loss of viability and also permitted the destruction of the DNA in the presence of DNase 1. This could explain the low yields using the commercial QIAGEN kit as it has both PEG precipitation and DNase 1 steps. Ultracentrifugation, in which the bacteriophage is pelleted by gravity, was shown to be superior to methods involving the addition of reagents for precipitation in this work.

Ultracentrifugation was also preferred as the reagents used in other methods have the potential for interfering with downstream applications. It has been shown that the ability of bacteriophage to attach to the host can be strongly affected by the ionic composition of the adsorption media (Hayes, 1964). This may have played a role in the low infectivity in several of the methods. With respect to the magnesium floc method, Vilagines *et al.,* (1982) has indicated that virions are non-infective when magnesium floc is formed *in situ*, but not when preformed, as was the case here. However, this previous work was carried out using polio virus and bacteriophage may be more sensitive to this protocol. Previous work found the PEG method to be destructive to bacteriophage in the family *Myoviridae* (Oakey and Owens, 2000) indicating that this is not an acceptable method when infection with the concentrate is required.

Extraction of bacteriophage DNA using a kit was found to result in low yields, with isopropanol concentration steps being insufficient. This kit was also difficult to handle and a phenol precipitation was trialed. The phenol precipitation resulted in good DNA yields when DNase 1 was not used. DNase 1 was used to eliminate any bacterial DNA present prior to extraction of bacteriophage DNA from its capsid. DNase 1 lowered yields dramatically, probably due to damage to the DNA when mechanical damage to the bacteriophage permitted entry of the DNase 1 to the capsid. In a study by Strauch *et al.* (2001), low DNA yields were not noted using kits involving DNase 1. However, this bacteriophage was produced in high concentration and a low percent yield of DNA may not have been considered relevant. Similarly, in studies using phenol extraction with a DNase 1 step (Silhavey *et al.*, 1984; Ogunseitan *et al.*, 1992), sufficient bacteriophage DNA was extracted. Addition of DNase 1 prior to plaque assay with a bacteriophage (family *Myoviridae*) from *Salmonella typhimurium* (Mirold *et al.*, 1999) had no effect on titre, indicating DNase 1 did not enter the capsid in that case. This may mean that in the case of BupsΦ1 there was excessive mechanical damage, due either to rough handling or an unusual level of fragility of the bacteriophage structure.

Previously, the prime use for restriction digestion of bacteriophage DNA has been for restriction mapping of the bacteriophage genome (Strauch *et al.*, 2001). However, restriction digest patterns have also been used to differentiate between bacteriophage (Ogunseitan *et al.*, 1992). Closely related bacteriophage, one being a mutant derived from the other, can be differentiated using this technique (Gertman *et al.*, 1986). Restriction digestion could also be used to determine whether infection was successful and whether any contamination with another bacteriophage had occurred. This can be a problem when serial amplification is undertaken (Billing, 1969). A variety of restriction enzymes have been used with the selection of enzyme dependant on the frequency of cut sites in each case. *Eco*RV has previously been used with the ΦCTX bacteriophage from *Pseudomonas aeruginosa* and in that situation, cut the genome into six fragments (Hayashi *et al.*, 1993). When bacteriophage from NCTC

13178 were cut with this enzyme, eight segments, six between one kb and ten kb and two greater than ten kb were produced.

Only those isolates that tested negative for bacteriophage by all assays were considered as potential hosts for bacteriophage quantification and amplification. As the first potential host tested (#4) was successful, the other candidates were not examined further. The low number of potential hosts indicates that infection of *B. pseudomallei* with bacteriophage is commonplace. Previous studies have shown high infection rates of *B. pseudomallei* with bacteriophage (Denisov and Kapliev, 1991; Denisov and Kapliev, 1995), with one study indicating a 91% infection rate of 35 isolates (Manzeniuk *et al.*, 1994). It is expected that if further isolates where tested, more potential hosts would be identified, but the percentage of suitable hosts would remain low.

4.4.4 Increasing titre

The high titre amplification of a temperate bacteriophage can be problematical, as a proportion of those cells infected will lysogenise rather than lyse (Billing, 1969). To minimise this, conditions for lysis must be optimal. The choice of host, inoculating dose of bacteriophage and other physiological conditions can affect whether lysis or lysogeny will occur (Jacob and Wollman, 1959). In this study, one host was found to be suitable, resulting in the #4(BupsΦ1) system. The fact that an inducer such as mitomycin C was not required to trigger mass lysis in this host-phage system indicates that the system is not as lysogenically stable as the original system from which the bacteriophage was extracted. Under certain conditions, complete lysis was possible, resulting in high bacteriophage yields.

In addition to the mitomycin C assay, there are other methods of induction in lysogenic systems. Bacteriophage lambda production can be controlled by temperature changes (Young, 1992), CS112 production from group A streptococci can be induced with a soluble product, SPIF, released by pharyngeal cells (Broudy *et al.*, 2001) and the stress of entering the stationary phase after logarithmic growth can

also trigger bacteriophage production. The latter type of induction may have played a role in the production of bacteriophage in this study, with low levels of production and infection occurring until an appropriate cell density was reached and mass lysis was triggered. This could explain the delay period between infection and complete lysis. It should be noted that this mass lysis does not occur when the original strain from which the bacteriophage is isolated is inoculated to high cell densities.

In lytic systems, inoculation with a ratio of 1:1 to 10:1 pfu:cfu has been shown to result in the greatest yield of bacteriophage. This is not the case in lysogenic systems. When the multiplicity of infection is small, 1:1 or less, most of the infected bacteria lyse. As the multiplicity of infection increases, the proportion of infected cells which lysogenise also increases (Boyd, 1953; Lieb, 1953), indicating that in some way the infecting particles can cooperate to establish lysogeny. In this study, when a low inoculum was used, analysis of free bacteriophage in supernatant in the first two hours showed low yields (Figure 4.5). Lower yields may occur in the early stages due to continual infection and lysis cycles in the broth, until yield is sufficient to cause visible lysis. It may be related to physiological conditions in the broth, with increasing growth and dropping levels of nutrition triggering lysis more frequently. As inoculation dose rose to one pfu/cfu, yield also rose, and as inoculation rose further, there was either little increase, or slight drop in yield. Slight variations in yield between experiments may be due to some physiological factor caused by different volumes in different containers, which itself affected the lysis:lysogeny ratio.

In the #4(BupsΦ1) system, the inoculating dose was optimised for maximum yield, but optimisation of physiological conditions was limited. One of the conditions that can alter the lyis:lysogeny ratio is temperature, with lower temperatures causing a swing towards lysogenesis (Bertani and Nice, 1954) and delaying bacteriophage reproduction (Fischetti *et al.*, 1971). This is not a universal effect, with P2 phage not affected by temperature alterations. In the current study, inoculation at low temperatures (4°C) followed by incubation at 37°C resulted in lysis being delayed

relative to inoculation at 37°C. This delay may have been caused by an initial tilt in the balance towards lysogeny. However, after an increase in temperature, lysis proceeded, indicating that any lysogeny was not highly stable. Other physiological conditions, such as bacterial nutrition levels can also effect the ability of bacteria to produce bacteriophage. Previously, under starvation conditions, bacteriophage production could not be induced (Jacob and Wollman, 1959). Enrichment of nutrient conditions may be of use to further optimise conditions in this study.

Once infected, lysis in a temperate system generally takes longer than with a lytic (or virulent) bacteriophage (Jacob and Wollman, 1959). In lysogenic systems, harvesting of bacteriophage typically takes place between two and six hours post-induction (Billing, 1969; Kidambi *et al.*, 1996; Williams *et al.*, 2002a), though lysis times of less than one hour have been recorded (Ptashne *et al.*, 1982; Young, 1992). Bacteriophage are generally collected from lysogenic systems around two hours after induction (Billing, 1969; Kidambi *et al.*, 1996; Williams *et al.*, 2002a) while lytic bacteriophage can reproduce and lyse the bacterial host in under one hour (Douglas, 1975; Ackermann, 1998). In the #4(BupsΦ1) system, at a dose of one pfu/cfu, lysis took seven hours. This is rather long, even for a temperate system, however no specific induction was undertaken. The broth used in this experiment was chilled, and cold temperatures are known to slow bacteriophage replication as noted above. However, the media was warmed to 37°C within 15 minutes so lysis should only be delayed by that amount of time unless the cold temperature causes the bacteria to enter lag phase, in which case recovery could be longer. There may be some lysogenisation occurring, followed by induction of lysis by a physiological factor such as cell density. An alternate explanation may be that several cycles of lysis may be occurring, with released bacteriophage infecting previously uninfected cells. At infection rates of less than one pfu/cfu, this can explain the increasing bacteriophage yield observed over time (Figure 4.5).

Of the six bacteriophage extracts which were analysed by T.E.M., five were found to belong to the family *Myoviridae* and one (#7) could not be classified. Previous work

with spontaneously lytic bacteriophage from *B. pseudomallei* have found family *Myoviridae* or *Siphoviridae* or, in some cases, both types present in the same bacterial isolate (Denisov and Kapliev, 1995; Woods *et al.*, 2002). It is likely that if enough of the bacteriophage extracted in this study were analysed by T.E.M., bacteriophage of the family *Siphoviridae* would also be found.

4.4.5 Characterisation of the bacteriophage

Pulse field gel electrophoresis (PFGE) has been used to determine the size of bacterial genomes, by restriction digestion and addition of the sizes of the fragments (Ratnaningsih *et al.*, 1990). This fragmentation is necessary to linearise the genome and limit the length of any individual fragment of DNA such that the lengths are measurable by gel electrophoresis. The creation of very small fragments along with those in the right size range for measurement may result in underestimation of genome size. Similarly, bacteriophage have been sized using restriction digestion and analysis via standard gel electrophoresis (Yoon *et al.*, 2001). Bacteriophage are typically already in an ideal size range for PFGE and as such, as long as they are shown to be linear, can be directly sized on PFGE without digestion. PFGE provides an effective means to size bacteriophage genomes without potential underestimation of size brought about by loss of small fragments. The genome of BupsΦ1 was found to be 55.1kb by this method.

Bacteriophage double stranded DNA genomes are typically greater than 20kb in length (Bernhardt *et al.*, 2002). Temperate bacteriophage in the family *Myoviridae* often have genomes smaller than that of BupsΦ1. For example; HP2; 31.5kb (Williams *et al.*, 2002a), VHML; 43.2kb (Oakey *et al.*, 2002), ΦCTX; 35.5kb (Hayashi *et al.*, 1990), sfV 37.1kb (Allison *et al.*, 2002), EJ-1; 42kb (Diaz *et al.*, 1992). There are also bacteriophage with larger genomes ; Φ7888; 63.4kb (Strauch *et al.*, 2001). As such, it is evident that bacteriophage BupsΦ1 fits within the expected genomic size range of the family *Myoviridae*.

4.4.6 Conclusion

The methods described above have been used to examine isolates of *B. pseudomallei* for the presence of bacteriophage and determine a gold standard for this purpose. Methods for extraction and amplification of bacteriophage from *B. pseudomallei* as well as DNA extraction from bacteriophage were optimised. Extraction and purification of a bacteriophage from *B. pseudomallei* NCTC 13178 was carried out and it was characterised as a member of the family *Myoviridae* by T.E.M., and analysis of the nucleic acid structure.

CHAPTER 5

INFECTION OF LOW VIRULENCE ISOLATES OF *Burkholderia pseudomallei* WITH BACTERIOPHAGE EXTRACTED FROM HIGH VIRULENCE ISOLATES

5.1 Introduction

Horizontal gene transfer is an important method by which bacteria alter their genetic information and increase their fitness (Miao and Miller, 1999). This has implications for developing new strains with epidemic potential (Faruque *et al.*, 1998). Horizontal gene transfer can occur by several means with conjugation, transformation and transduction being the most common of these (Miller, 1998). Conjugation is the transfer of DNA from one live bacterium to another via a pilus or chemical signalling drawing the bacteria together. Transformation is the uptake of free DNA, in the form of plasmids or DNA released from dead bacteria closely related to the bacteria that take it up. In transduction, bacteriophage pick up genetic material from one bacterial cell and deposit it in another. Bacteriophage are capable of transferring whole plasmids and pieces of chromosomes between hosts (Miller, 1998; Weinbauer and Rassoulzadegan, 2004).

One limit to the transfer of genetic information by transduction is the existence of superinfection immunity. This form of resistance to infection with bacteriophage is caused by temperate bacteriophage already present in the bacteria stopping further infection and lysis by similar or the same bacteriophage (Bertani, 1971). There are multiple mechanisms for superinfection immunity (McGrath *et al.*, 2002). O-antigen modification can alter the polysaccharide backbone structure in the cell wall, thus changing receptor structure such that bacteriophage can no longer attach to the cell wall (Allison *et al.*, 2002). A repressor mediated lambdoid immunity mechanism exists which prevents transcription of lytic genes (Strauch *et al.*, 2004) and a transcription termination mechanism, by which expression of genes involved in the lytic cycle is inhibited by RNA-RNA interactions, has also been identified (Allison *et al.*, 2002). The presence or absence of superinfection immunity is one of the

methods by which researchers have analysed and grouped bacteriophage (Wagner *et al.*, 1999; Strauch *et al.*, 2004).

When in the lysogenic part of their lifecycle, the temperate bacteriophage genome is inserted into the host chromosome. Insertion may be to a single site on the bacterial chromosome or to any one of a number of sites (Casjens, 2003). A common site of bacteriophage insertion into bacterial chromosome are the tRNA genes on the bacterial chromosome (Cheetham and Katz, 1995). As these are present in multiple copies, they can provide multiple potential integration sites. Once inserted, genes on the bacteriophage can be expressed, conferring new metabolic traits on the host. The process of altering properties of host bacterial cell on establishment of lysogeny is known as phage conversion. (Acheson *et al.*, 1998). This can increase the fitness of the host in its environment and thus aid survival of the bacteriophage (Miao and Miller, 1999). This has particular relevance in the field of emerging infectious diseases. Enterohaemorrhagic *E. coli* (*EHEC*) is believed to have arisen from an enteropathogenic *E.coli* progenitor (*EPEC*) which acquired prophage encoded Shiga-like toxins, thus becoming a new pathogen that expressed both sets of traits (Whittam *et al.*, 1993).

During replication in the bacterial cell, bacteriophage DNA is copied from the bacterial chromosome for insertion into the progeny bacteriophage capsid. If bacterial chromosomal information is mistakenly copied, it can be transferred to the next bacteria infected by the progeny bacteriophage. In many cases, where virulence genes are carried on bacteriophage, they are located near the bacteriophage attachment site and integrase gene, indicating they have been acquired by transduction (Cheetham and Katz, 1995).

Virulence factors which are encoded on bacteriophage can form various functions. Those identified so far include genes expressing extracellular toxins, antigenic alteration proteins, effector proteins involved in invasion, enzymes required for intracellular survival, serum resistance from outer membrane proteins and adhesins

for bacterial host attachment as well as a range of other functions (Bensing *et al.*, 2001; Boyd and Brussow, 2002).

Several bacterial - bacteriophage systems have been analysed in which bacteriophage extracted from a virulent strain of bacteria can be inserted into a strain of lower virulence and can cause an upregulation of virulence. Bacteriophage VHML (Oakey and Owens, 2000), extracted from a virulent strain of *Vibrio harveyi* and transferred to a strain of lower virulence, was shown to upregulate virulence in salmon species, *Artemia* (Austin *et al.*, 2003) and *Penaeus monodon* (Munro *et al.*, 2003). It also caused enhanced haemolytic activity on trout blood agar plates and sheep blood agar plates and proteins excreted from converted *Vibrio harveyi* were identified as exotoxins by monoclonal antibody assays. Sequencing of the VHML genome also identified a putative transcriptional regulator, N6-Dam, a methyltransferase which may activate or repress host genes (Oakey *et al.*, 2002).

Many experiments to determine whether bacteriophage can infect uninfected strains of bacteria have been carried out *in vitro*, mimicking transduction in the environment. However, it is also possible for factors in a mammalian host to be necessary for successful transduction to occur (Mel and Mekalanos, 1996). It has been determined that infectious *V. cholerae in vivo* assays cause significant enrichment of toxigenic *V. cholerae* (Baselski *et al.*, 1979) and that the bacteriophage present in toxigenic *V. cholerae*, CTXΦ, can be produced from lysogens present in the intestine (Lazar and Waldor, 1998). The diarrhoea caused by the toxin present on this bacteriophage can aid in dissemination of *V. cholerae* and hence, survival. In colonisation of the small intestine by *V. cholerae*, there is evidence that bacterial expression of toxin co-regulated pili (TCP) is triggered by host environmental signals and that these TCP are the receptor for CTXΦ, the bacteriophage encoding the cholera toxin. *In vivo* lysogenic conversion of an El Tor strain occurs 10^5-10^6 times more efficiently than *in vitro* conversion (Waldor and Mekalanos, 1996).

Of interest in this study is the effect of infection of isolates *of B. pseudomallei* with bacteriophage extracted from other isolates of *B. pseudomallei*. *B. pseudomallei* is a human pathogen with a broad virulence range (Ulett *et al.*, 2001). The genome of strain K96243 of *B. pseudomallei* has been sequenced and determined to consist of two circular replicons (European Molecular Biology Laboratory accession numbers BX571965 and BX571966) of 4.07Mb (chromosome 1) and 3.17 Mb (chromosome 2). Analysis of these chromosomes indicates the presence of multiple, recently acquired genetic islands (GIs) with at least three appearing to be prophages. In addition, comparison of GIs with other strains indicated their presence to be highly variable between strains (Holden *et al.*, 2004).

Many virulence factors of *B. pseudomallei* have been identified, including a type three secretion system (TTSS) (Winstanley *et al.*, 1999), an exotoxin (Mohamed *et al.*, 1989a) and flagella (Chua *et al.*, 2003), with potential virulence determinants including surface polysaccharides, fimbrae, pili, drug resistance determinants and putative adhesins (Holden *et al.*, 2004).

Given the range of toxins and other virulence determinants found on bacteriophage, and the presence of some of these virulence determinants in *B. pseudomallei*, it is possible that lysogenic bacteriophage carried on *B. pseudomallei* of high virulence may be harbouring some bacterial virulence determinants. Isolates of low virulence may be lacking such genes. If so, insertion of bacteriophage from an isolate of high virulence to one of low virulence may cause upregulation of virulence, as has been observed in other bacterial systems.

The aim of this chapter is to examine whether bacteriophage extracted from bacterial isolates of high virulence, as determined in the Balb/C mouse model (Leakey *et al.*, 1998), can infect isolates of lower virulence and alter physiological factors including virulence.

Isolates of interest include the highly virulent isolate NCTC 13178 and its isogenic mutant of low virulence NAFC. While previous work (Section 4.3.1.5) has shown that NAFC still contains a bacteriophage with an identical restriction endonuclease cutting pattern and should therefore display superinfection immunity to the bacteriophage BupsΦ1 isolated from NCTC 13178, cross spotting patterns of the two isolates are very different. It is possible that the process of producing NAFC has caused a change in the prophage which is not identifiable by restriction digest assay. It is also possible that superinfection immunity, as well as virulence, may have been lost.

Other *B. pseudomallei* isolates, which have been found by all methods tested (Table 4.1) to be negative for inducable bacteriophage, will also be examined with respect to infection with BupsΦ1. If attenuation of virulence in NAFC is not caused by alterations to its bacteriophage, addition of BupsΦ1 is unlikely to increase virulence. However BupsΦ1 may still carry virulence genes and examining other bacterial isolates not already carrying a copy of this genome could determine this. It will also be possible to determine successful insertion of the BupsΦ1 genome into bacteria not already containing a copy of the genome. One other bacteriophage extracted from a bacterial isolate of high virulence will also be examined. As bacteriophage Φ7 has a different restriction digest pattern to BupsΦ1, successful insertion into NAFC will be detectible and this bacteriophage can be used as a control for the experiment.

5.2 Materials and Methods

5.2.1 Infection with bacteriophage

Parent bacteria were cultured on Ashdown agar for 48 hours and a single colony inoculated into LB broth. The broth was incubated at 37°C in an orbital shaker at 100rpm until an $O.D._{600nm}$ of approximately 0.1 was reached. The broth was then spiked with bacteriophage previously extracted (from donor bacteria) at a dose of 100µl of bacteriophage solution (at about $5x10^9$ pfu/ml) per 100ml of broth. The spiked broth (offspring bacteria) was incubated overnight at 37°C in an orbital shaker at 100rpm and a sample was stored in 10% glycerol (v/v) and frozen. A loop of the

broth was also plated on sheep blood agar (SBA) and Ashdown agar, incubated at 37°C and observed at 24 hour intervals.

A sample of bacteriophage was tested in each case for contamination with bacteria by spiking ten µl of sample into ten ml of sterile LB media and incubating at 37°C with shaking at 100rpm for 48 hours. Bacteriophage was considered uncontaminated with bacteria if no opacity was observed in the media after this time.

5.2.2 Colony morphology

Donor, parent and offspring bacteria, as described above, were examined and compared. A loop of bacterial broth was cultured on Ashdown agar, incubated at 37°C and examined after three to five days. Colonial morphology was observed and described. Photographic comparison was undertaken using either 35mm film or digital photography.

5.2.3 Selection of offspring bacteria

Two methods of selection for inoculation in mice were tested. In the first method, those colonies of offspring bacteria which did not show similar morphology on Ashdown agar to the parent line they were created from, were preferentially selected. They were then inoculated into ten ml LB broth in parallel with a colony of parent bacteria. These were incubated at overnight 37°C with shaking at 100rpm. A one ml aliquot of the offspring bacteria in 10% glycerol (v/v) was frozen at -70°C. Confirmation of infection was determined by extraction of bacteriophage from the offspring sample (Section 4.2.2.1), followed by DNA extraction and restriction digest assay (Sections 4.2.3 and 4.2.3.1). The bacteriophage was considered to be present when the restriction digest pattern included the bands which were present when the donor bacteriophage was digested.

In the second method, broth spiked with bacteriophage the previous day was used directly for mouse ID_{50} assays (Section 5.2.4) without an intermediate step involving growth on agar and selection of a single colony.

5.2.4 Mouse ID_{50} assay

Using the first method (Section 5.2.3), bacteria were cultured on Ashdown agar for 48 hours and a single colony inoculated into ten ml LB broth (unless noted otherwise in Figure 5.1). The broth was incubated at 37°C in an orbital shaker at 100rpm until an $O.D._{600nm}$ of approximately 0.2 was reached. In both methods, ten fold serial dilutions were then carried out and the bacteria in these dilutions was quantified using a Miles-Misra assay (Section 3.1.3). Five replicates of 200µl each of the serial suspensions were injected into Balb/C mice via the tail vein. A range of dilutions were selected, dependant on the ID_{50} of the parent bacteria and the ID_{50} of the donor bacteria. Mice were examined at least twice daily for seven days for evidence of illness and moribund mice were removed and euthanased. At the end of seven days, surviving mice were also euthanased. Autopsies were carried out to examine spleens and observations were recorded. The ID_{50} was determined using the protocol described in Section 3.3.1 and Appendix 4.

Passaging of isolates was carried out by homogenising spleens extracted from moribund mice in a stomacher (Stomacher Lab Blender 80; Townson and Mercer, Australia) in five ml PBS. The homogenate was either used directly for a second ID_{50} assay or plated on Ashdown agar and a resultant colony inoculated into LB as per Section 5.2.3.

Five isolates of *B. pseudomallei* (NAFC, E4, #69, #83, #13) were infected with bacteriophage from *B. pseudomallei* sources. All isolates were infected with BupsΦ1 (from NCTC 13178) and isolates NAFC, #13, #69 and #83 were infected with Φ7 (from #7). Four of the isolates were selected as being negative for the presence of bacteriophage using the gold standard designed in Chapter 4. In addition to this, three of these (E4, #69, #83) had an ID_{50} of greater than $1x10^4$ cfu in the Balb/C mouse model. Isolate #13 had an ID_{50} just below 10^4 cfu ($3.8x10^3$ cfu). The fifth isolate (NAFC), determined to be positive for bacteriophage in Chapter 4, was the low virulence isotype of NCTC 13178. The bacteriophage *Eco*RV digest pattern was determined to be similar to that of the bacteriophage of NCTC 13178 (Section

4.3.1.5). This isolate was used to determine whether the loss of virulence could be restored by addition of additional bacteriophage extracted from the donor isolate.

Flow charts of the experiments undertaken can be seen in Figure 5.1. Each infection of bacteriophage was given a letter (A-H) as an identifier and any assay carried out was given a subcode, eg., A.1, A.1.1, A.2. BupsΦ1 is described as Φ1 for brevity. Each separate infection of bacteria with bacteriophage is highlighted in blue. Each isolate for which an ID_{50} was determined is noted and highlighted in yellow. One assay (Assay C.1) used bacteriophage concentrated by the magnesium hydroxide assay (4.2.2.2) instead of the standard ultracentrifugation method (4.2.2.1).

Infection with bacteriophage was presumed to have occurred if colony morphology of the offspring bacteria changed from that of the parent isolate. Confirmation was carried out by identification of restriction digest pattern of the bacteriophage where possible.

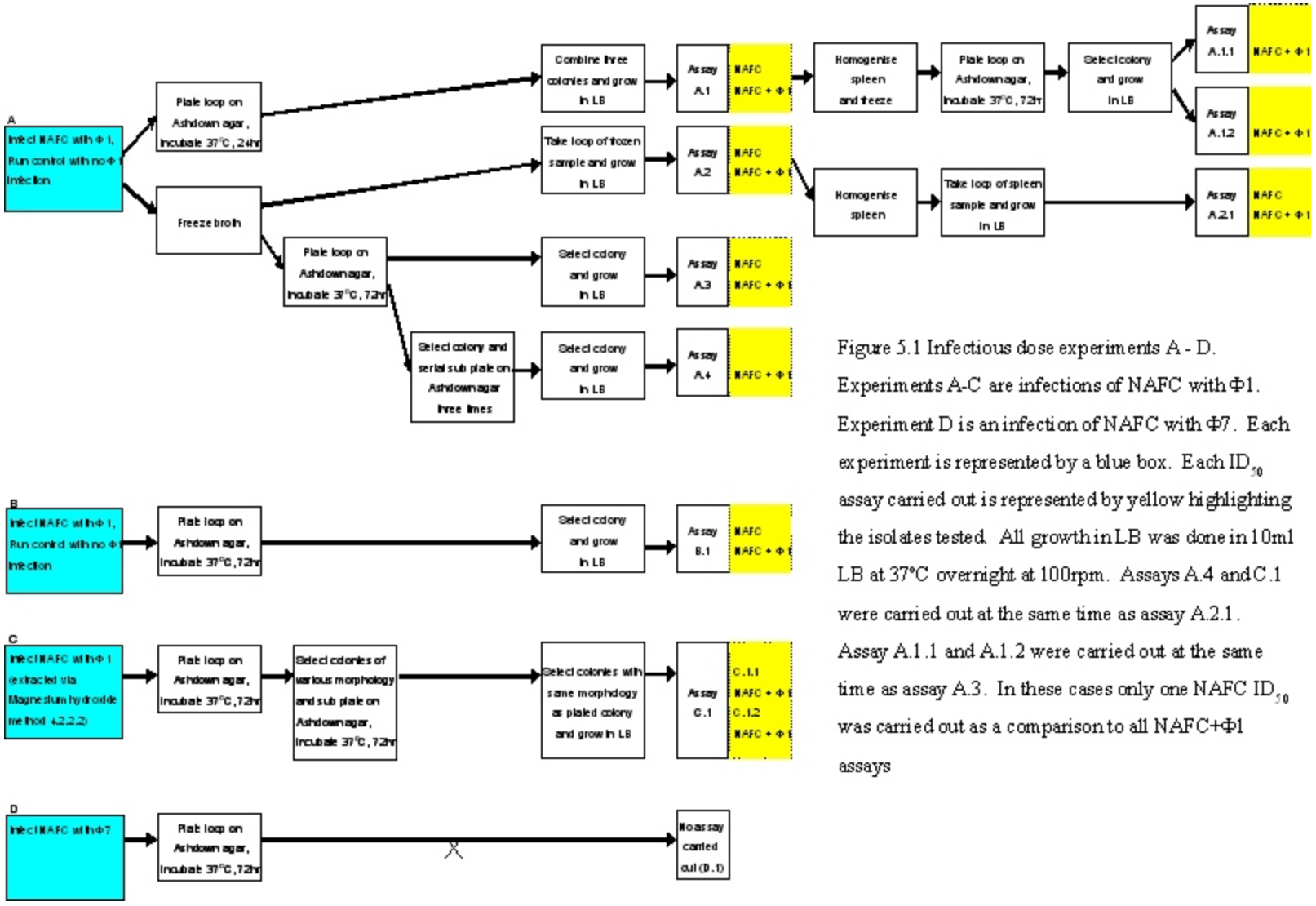
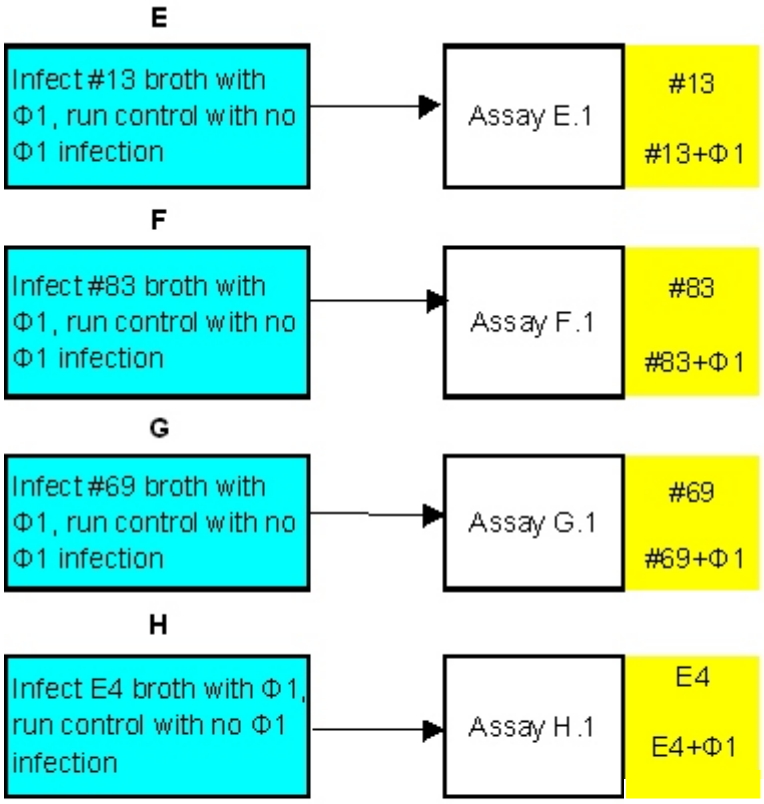


Figure 5.1 Infectious dose experiments A - D. Experiments A-C are infections of NAFC with Φ1. Experiment D is an infection of NAFC with Φ7. Each experiment is represented by a blue box. Each ID_{50} assay carried out is represented by yellow highlighting the isolates tested. All growth in LB was done in 10ml LB at 37°C overnight at 100rpm. Assays A.4 and C.1 were carried out at the same time as assay A.2.1. Assay A.1.1 and A.1.2 were carried out at the same time as assay A.3. In these cases only one NAFC ID_{50} was carried out as a comparison to all NAFC+Φ1 assays



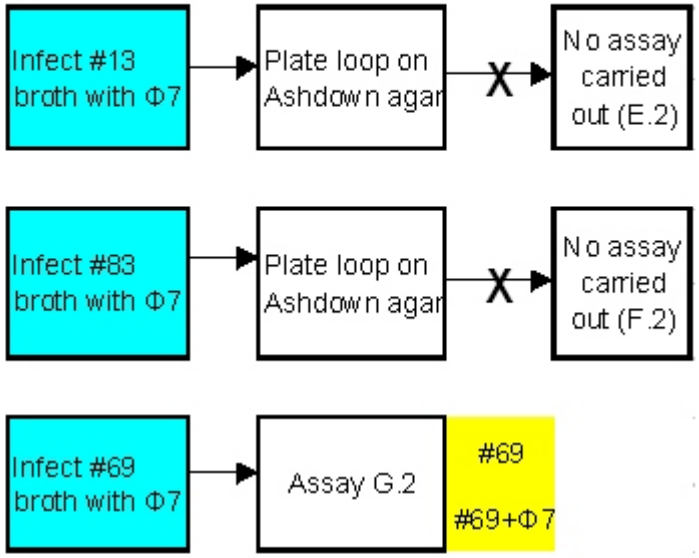


Figure 5.1 E-H. Infection dose experiments of isolates other than NAFC with either Φ1 or Φ7. Each experiment is represented by a blue box. Each ID_{50} assay carried out is represented by yellow highlighting the isolates tested. All growth in LB was done in 10ml LB at 37°C overnight at 100rpm

5.2.5 Plaque observation on lawns

Bacteria were cultured on Ashdown agar for 48 to 72 hours and a single colony inoculated into LB broth. The broth was incubated overnight at 37°C in an orbital shaker at 100rpm. To eliminate the possibility that plaques were present but difficult to see due to low bacterial growth, volumes of overnight broths added to top agar and volumes of top agar used were varied. One hundred (Group 3) to two hundred microlitres (Group 2) of broth was added to ten ml of molten soft top agar, with or without addition of mitomycin C to a concentration of 100mg/ml and then poured over LB agar pre-warmed to 37°C. Alternately, broth was incubated to an $O.D._{600nm}$ of 0.5 and 200µl was added to five ml of molten soft top agar, prior to pouring over pre-warmed LB agar (Group1). Group one and two plaque tests used broths inoculated from 48 hour old Ashdown colonies. Group three plaque tests used broths inoculated from 72 hour Ashdown colonies. In each case the agar plates were allowed to solidify and then inverted and incubated at 37°C overnight. Plates were then examined for the presence of plaques.

5.3 Results

5.3.1 Infectious dose assays

Photographic plates of colony morphology of bacterial isolates from which bacteriophage was extracted for these assays (donor isolates) demonstrate both isolates as being rough and flat with rugose to umbonate edges (Figure 5.2). Photographic plates of an example of colony morphology of a donor, parent and offspring bacteria are in Figure 5.3. In this example, clear changes in colony morphology were observed with offspring bacteria more closely resembling the donor isolate than the parent isolate. Plates of other colonial morphology can be found in Appendix 7. In most experiments, plating broth to which bacteriophage had been added resulted in at least some colonies retaining parental morphology. Colonies which showed morphology that was altered from that of the parent were selected and sub-plated for the images in Appendix 7. Colony morphology is described in Table 5.1. This table also contains results of infectious dose experiments where these were carried out.

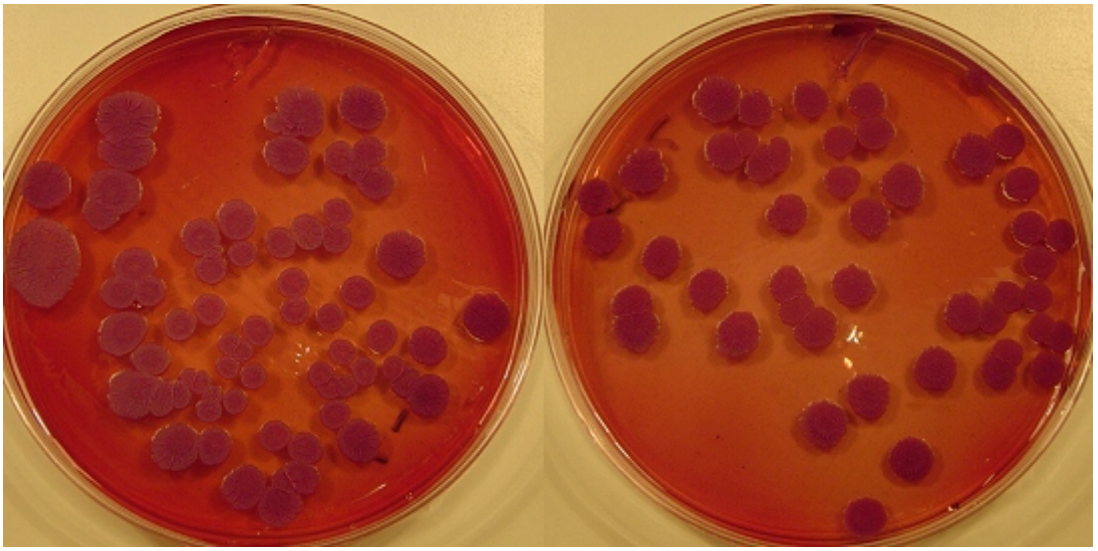


Figure 5.2 Ashdown plates of donor *Burkholderia* isolates. Plates contain from left to right; NCTC 13178, and isolate #7 after five days of incubation at 37°C on Ashdown agar. Donor isolate morphology can be described as follows; *B. pseudomallei* NCTC 13178: five to ten mm, flat, rhizoid to radially striated with a lobate edge, dull, rough, purple. *B. pseudomallei* #7: seven mm, flat, rhizoid, umbonate, dull, rough, purple

Figure 5.3 Example of colony morphologies of donor, uninfected parent and offspring bacteria on Ashdown agar. Plates from left to right are; Donor: NCTC 13178, Parent: NAFC, Offspring: NAFC-BupsΦ1 after five days incubation at 37°C on Ashdown agar

Table 5.1 Results of virulence assays including ID_{50} data in mice for both parent isolates and offspring isolates, as well as observations of colonial morphology and restriction digest results. The ID_{50} of the donor isolates for these assays were; NCTC 13178 (BupsΦ1) 9 colony forming units (cfu), #7 (Φ7) 40 cfu (Ulett *et al.*, 2001).

Assay	Isolate	Level of confirmation of infection reached	ID_{50} (cfu)	Colonial morphology on Ashdown agar after five days incubation
A.1	NAFC	NAFC/BupsΦ1 digest pattern detected	$2.5x10^4$	3-10mm, convex, entire, glistening, smooth mucoid, lilac
A.1	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$8.2x10^1$	4-10mm, flat, undulate to entire, rhizoid, glistening, rough, lilac
A.1.1	10^3 spleen homogenate	BupsΦ1 digest pattern detected	$1.7x10^4$ (NAFC control A.3 $2.5x10^4$)	similar to assay A.1 NAFC+BupsΦ1
A.1.2	10^1 spleen homogenate	BupsΦ1 digest pattern detected	$1.0x10^6$ (NAFC control A.3 $2.5x10^4$)	similar to assay A.1 NAFC+BupsΦ1
A.2	NAFC	NAFC/BupsΦ1 digest pattern detected	$2.2x10^4$	similar to assay A.1 NAFC
A.2	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$1.6x10^5$	4-6mm flat to convex, entire, glistening, smooth, mucoid, purple (darker and flatter than NAFC)
A.2.1	NAFC	NAFC/BupsΦ1 digest pattern detected	$1.8x10^5$	similar to assay A.1 NAFC
A.2.1	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$2.5x10^6$	similar to assay A.1 NAFC+BupsΦ1
A.3	NAFC	NAFC/BupsΦ1 digest pattern detected	$2.5x10^4$	similar to assay A.1 NAFC
A.3	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$4.1x10^4$	similar to assay A.1 NAFC+BupsΦ1
A.4	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$3.6x10^5$ (NAFC control A.2.1 $1.8x10^5$)	similar to assay A.1 NAFC+BupsΦ1
B.1	NAFC	NAFCBupsΦ1 digest pattern detected	$6.5x10^5$	similar to assay A.1 NAFC
B.1	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$5.5x10^5$	similar to assay A.1 NAFC+BupsΦ1
C.1.1	NAFC+ BupsΦ1	BupsΦ1 digest pattern detected	$4x10^4$ $1.3x10^5$ (NAFC control A.2.1 $1.8x10^5$)	2-7mm, convex, entire, glistening, smooth, mucoid, purple
C.1.2	NAFC+ BupsΦ1	no digest pattern detected	$2.3x10^6$ (NAFC control A.2.1 $1.8x10^5$)	similar to assay A.1 NAFC
D.1	NAFC+ Φ7	BupsΦ1 digest pattern detected, Φ7 digest pattern not detected	not done	similar to assay A.1 NAFC
E.1	#13	no digest pattern detected	$3.8x10^3$	5mm, flat, undulate, rhizoid, dull, rough, purple
E.1	#13+ BupsΦ1	no digest pattern detected	$1.7x10^4$	10-12mm, flat, rhizoid, undulate to lobate, dull, rough, purple
E.2	#13+ Φ7	no digest pattern detected	not done	similar to assay E.1 #13
F.1	#83	no digest pattern detected	$1.68x10^4$	3-4mm, flat, entire, rhizoid, dull, rough, purple
F.1	#83+ BupsΦ1	BupsΦ1 digest pattern detected	$7.9x10^3$	4-5mm, flat, entire, rhizoid, dull, rough, purple

Table 5.1 continued

Assay	Isolate	Level of confirmation of infection reached	ID_{50} (cfu)	Colonial morphology on Ashdown agar after five days incubation
F.2	#83+ Φ7	no digest pattern detected	not done	similar to assay F.1 #83
G.1	#69	no digest pattern detected	$1.96x10^4$ - $8x10^4$	3-5mm, flat, rhizoid, entire, dull, rough, purple
G.1	#69+ BupsΦ1	BupsΦ1 digest pattern detected	$1.3x10^4$	3-5mm, flat – thinner than parent it was created from, rhizoid, undulate, dull, rough, translucent lilac
G.2	#69+ Φ7	Φ7 digest pattern detected	$7.4x10^4$	5-7mm, flat –thinner than parent it was created from, rhizoid, undulate, dull, rough, purple
H.1	E4	no digest pattern detected	$2.17x10^6$	5mm, convex, entire, glistening, smooth mucoid, lilac
H.1	E4+ BupsΦ1	no digest pattern detected	$1.2x10^6$	6mm, convex, undulate, glistening, smooth mucoid, lilac

Across all NAFC experiments, the NAFC ID_{50} varied from $2.2x10^4$ (assay A.2) to $6.5x10^5$ (assay B.1) cfu. The average NAFC ID_{50} from five assays was $1.8x10^5$ cfu with a standard error of $1.2x10^5$ cfu. Based on the variability in NAFC ID_{50}s, possibly caused by variability in batches of mice, standard error cannot be used to determine significant alterations in virulence. Instead, an alteration in virulence from NAFC control to NAFC+BupsΦ1 experiment was considered to be significant if it exceeded one log in value.

The initial trial infecting NAFC with BupsΦ1 resulted in an upregulation of virulence (Table 5.1, assay A.1) with the ID_{50} of the offspring ($8.2x10^1$ cfu) approaching that of the donor isolate (9 cfu).

Two moribund mice from assay A.1, one at 10^1 cfu dose (NAFC+BupsΦ1) and the other at 10^3 cfu dose, were autopsied and their spleens extracted. At extraction, the spleens were enlarged with one mm diameter white pustules and two mm diameter dark discolourations. After five days on Ashdown agar, colonies extracted from the spleen of the mouse dosed at 10^3 cfu were three mm diameter, shiny, low convex, purple colonies with some flattening. Plates from the mouse spleen dosed at 10^1 cfu displayed a mixed morphology; some were shiny, smooth, 15-20 mm diameter colonies and others were 15mm diameter, flat, rhizoid colonies. The homogenates

were frozen at -80°C. Passaging homogenates, using colonies plated on Ashdown agar, resulted in ID_{50}s similar (Table 5.1, assay A.1.1, 10^3 spleen; $1.7x10^4$) to those of the uninfected parent NAFC ($2.5x10^4$; using NAFC from assay A.3 which was done in conjunction) and two logs higher (assay A.1.2, 10^1 spleen; $1x10^6$).

The other assays carried out from the first infection experiment after a variety of further steps were undertaken (assays A.2, A.3 and A.4; $1.6x10^5$, $4.1x10^4$, $3.6x10^5$ respectively) did not show an increase in virulence as was seen in the initial assay. Three spleens were extracted from the surviving mice from the direct inoculation method (assay A.2). These were processed for passage and no increase in virulence occurred during passage. It was noted that the ID_{50} increased to $2.5x10^6$ cfu (assay A.2.1).

Fresh infections of NAFC with BupsΦ1 were carried out and single three day old colonies from Ashdown agar were selected (assays B.1 and C.1). In all cases, no increase in virulence was observed. It should be noted that assay C.1 was carried out using bacteriophage concentrated using a magnesium protocol (4.2.2.2 magnesium hydroxide precipitation) rather than ultracentrifugation (4.2.2.2 ultracentrifugation). The ID_{50}s of the offspring bacteria in this case spanned more than one log in value. These assays were done in conjunction with assay A.2.1 and the only NAFC+BupsΦ1 ID_{50} which was more than one log different to NAFC ($1.8x10^5$) in this case was C.1.2, showing a slight loss of virulence ($2.3x10^6$).

Alternate *Burkholderia* isolates were also infected with BupsΦ1. In each case, no significant upregulation of virulence was recorded. ID_{50}s were carried out directly on overnight infected broths in an attempt to select for virulent offspring bacteria. One other bacteriophage (Φ7) was also used to infect *Burkholderia* isolates. One ID_{50} was carried out (assay G.2; #69+Φ7) directly on overnight broth and no alteration of virulence from that of the parent was observed.

Several assays did not show the presence of bacteriophage via restriction digest but did show an alteration in offspring colony morphology. In these cases, virulence assays were carried out with no increases in virulence noted. Where no change in morphology was noted and restriction digest assay did not confirm the presence of bacteriophage, no virulence assay was carried out to limit the use of mice (Assays D.1, E.2 and F.2).

5.3.2 Lawns

Previously, spontaneous plaque formation was observed with some *Burkholderia pseudomallei* isolates (Section 4.3.1). Several isolates, including donor, parent and offspring bacteria, were further examined for plaque formation using the soft top agar technique. This was done to determine whether the offspring bacteria spontaneously entered the lytic part of the lifecycle (Table 5.2). Addition of mitomycin C was carried out in some cases to attempt to stimulate lysis.

Assays for plaque formation were done in three batches at three different times using three different bacteria to agar ratios. Group one showed no spontaneous plaque formation after 200µl of low density bacterial broth was added to 5ml of agar. In both groups two and three, isolate NCTC 13178 did not produce plaques, even with addition of mitomycin C. This is the isolate from which bacteriophage was extracted for all of the offspring assays in this experiment. Also, in both groups two and three, NAFC produced plaques irrespective of addition of mitomycin C. Once bacteriophage from NCTC 13178 had been added to NAFC plaques were generally not observed, with the exception of one plaque in group two from an isolate of NAFC+BupsΦ1 selected from Assay B.1 (Table 5.1) and multiple plaques present from an isolate of NAFC+BupsΦ1 selected from Assay A.1 (Table 5.1) in group three.

Table 5.2 Results of production of plaques in soft top agar. Plaque production is indicated with 'y'. No plaque production is indicated with 'n'. In the case of single or very low level plaque production, the number of the plaques is also presented. "+MitC" represents addition of mitomycin C at 100mg/ml. "-MitC" represents no addition of mitomycin C. Results are grouped with respect to top agar conditions used.

group	isolate	Origin of infected isolate - Assay number from table 5.1	plaques		Top agar conditions
1	NCTC 13178		n		200µl of O.D.$_{600}$ 0.5 in 5ml 0.7% agar
1	NAFC		n		200µl of O.D.$_{600}$ 0.5 in 5ml 0.7% agar
1	NAFC+ BupsΦ1	A.1	n		200µl of O.D.$_{600}$ 0.5 in 5ml 0.7% agar
			-Mitc	+Mitc	
2	NCTC 13178		n	n	200µl of overnight broth + 10ml 0.7% agar
2	NAFC		y	y	200µl of overnight broth + 10ml 0.7% agar
2	NAFC+ BupsΦ1	A.1	n	n	200µl of overnight broth + 10ml 0.7% agar
2	NAFC+ BupsΦ1	B.1	y(1)	n	200µl of overnight broth + 10ml 0.7% agar
3	NCTC 13178		n	n	100µl of overnight broth + 10ml 0.7% agar
3	NAFC		y	y	100µl of overnight broth + 10ml 0.7% agar
3	NAFC+ BupsΦ1	A.1	y	y	100µl of overnight broth + 10ml 0.7% agar
3	NAFC+ BupsΦ1	B.1	n	n	100µl of overnight broth + 10ml 0.7% agar
3	NAFC+ BupsΦ1	B.1 (with uninfected parent morphology)	n	n	100µl of overnight broth + 10ml 0.7% agar

5.4 Discussion

The aim of this chapter was to examine whether bacteriophage extracted from bacterial isolates of high virulence could infect isolates of lower virulence and alter physiological factors including virulence.

The presence of a bacteriophage in NAFC, with the same restriction cutting pattern as that of BupsΦ1, prevented confirmation of infection by restriction digest assay when NAFC was exposed to BupsΦ1. However, when BupsΦ1 was added to NAFC and the broth plated, the morphology of the majority of colonies changed to a flatter,

undulate, rhizoid, rough form indicating some effect of the bacteriophage on the bacteria. One possibility is that the bacteriophage is causing lysis of sensitive bacteria, which are in the majority. This would permit only the rare phage resistant mutants (which have are likely to have a different gross morphology) to predominate. The altered morphology was similar between colonies and experiments indicating, if it was a mutant, it was clonal. Newly cloned organisms should have identical virulence properties with very narrow variation in an assay (Table 5.1). This was not the result in the case of NAFC+BupsΦ1 virulence experiments, so this is unlikely to be the cause.

In addition, alkalisation of the agar was also noted, although it did not appear to be as extreme as in the case of NAFC. Colony morphology more closely matched the ancestral NCTC 13178. Serial sub-plating resulted in no reversal in colony morphology to that of NAFC. NAFC+BupsΦ1 colonies were not identical to the NCTC 13178 in appearance. This would be expected if the only change between the donor and parent were the alteration of a bacteriophage, which was repaired by replacement of that bacteriophage at the same insertion site on the chromosome.

Other lysogenic bacteriophage, including P2 (Barreiro and Haggard-Ljungquist, 1992) and bacteriophage Mu (Hsu and Davidson, 1972), have been shown to have multiple insertion sites and this may be the case here. Changes in colony morphology after infection with bacteriophage have been observed in various systems including pf4 infection in biofilms of *P. aeruginosa* (Webb *et al.*, 2004), *tox*+ infection in *Clostridium novyi* (Eklund *et al.*, 1976) and small colony variants in *E. coli* (Kuo *et al.*, 2000). Some of these changes have been attributed to lysis of the colonies. Colony lysis may be playing a role in NAFC+BupsΦ1 in the case of infections where virulence was slightly lower than the virulence of NAFC (A.1.2, A.2.1, C.1.b). If the lysogenisation is less stable than in the original NCTC 13178 due to position of insertion or interaction with the bacteriophage already present, lysis could occur and bacterial growth would be retarded. Studies with insertion of P2 into strains of *E. coli* has determined that when a cryptic bacteriophage is present at the preferred

attachment site, P2 inserts at another attachment site (Barreiro and Haggard-Ljungquist, 1992) and when two P2 bacteriophage are inserted in tandem, the insertion is unstable and leads to frequent excision of the bacteriophage and lysis (Bertani, 1971). This loss of virulence is of interest for therapeutic use of bacteriophage to treat *Burkholderia* infections. However, the possibility of an increase in virulence would have to be eliminated for such a path to be worth pursuing.

If superinfection immunity mechanisms exist on BupsΦ1, it should not be able to infect NAFC as the bacteriophage already present in the genome should make it immune to similar or identical bacteriophage. However, cross spotting experiments (Chapter 4.3.1) showed variation in the spotting profiles of NCTC 13178 and NAFC. Addition of bacteriophage extracts to these bacteria resulted in few cases of plaque production with NCTC13178 (7/50), but a large number of cases of plaque production with NAFC (31/50) . This may be due to a greater variety of bacteriophage being able to infect NAFC, due to changes in cell wall structure or loss of some other superinfection immunity mechanism. Alternatively, it may be due to addition of bacteriophage extracts stimulating conversion of BupsΦ1 to a lytic lifecycle with NAFC. As plaques were not examined at the molecular level, this second possibility could be neither confirmed nor ruled out.

While it was often possible to alter colonial morphology by addition of bacteriophage, only one assay resulted in a significant increase of virulence (A.1). One possible reason for this is a requirement for the bacteriophage to insert at a particular site to increase virulence, possibly by insertionally inactivating a bacterial gene. Three assays (A.1.2, A.2.1, C.1.b) resulted in slightly lower virulence which adds evidence to the possibility of insertion position playing a role in the interaction of bacteriophage and bacteria. Insertion at different sites on the genome by bacteriophage Mu in *E.coli* has produced dissimilar phenotypes caused by gene inactivation at different insertion sites (Taylor, 1963; Hsu and Davidson, 1972) and these changes could affect the fitness and virulence of the bacteria.

Since repeating the NAFC experiment (assays B.1, C.1) did not result in the same increases as those of A.1 and cultures saved from this experiment did not cause increases in virulence (assays A.1.1, A.1.2, A.2, A.2.1, A.3), it cannot be definitively stated that BupsΦ1 increases virulence, only that there may be a link between virulence and the bacteriophage.

To examine whether BupsΦ1 was able to increase virulence in isolates of *B. pseudomallei* which did not show evidence of an inducible bacteriophage, several isolates of *B. pseudomallei* were infected. Colony morphology changes were noted and in the case of #69+BupsΦ1 (assay G.1) the changes were considered similar to changes noted with NAFC+BupsΦ1, in that colonies appeared thinner. However, in no case did an increase in virulence occur. Direct infection with overnight broths was trialed as some experiments have shown an *in vivo* effect with enrichment of virulent isolates (Baselski *et al.*, 1979; Waldor and Mekalanos, 1996), but in this case no enrichment of any virulent offspring occurred.

The alternate possibility is that BupsΦ1 does not encode virulence factors. Bacteriophage P1 confers no virulence changes on its host (Yarmolinsky, 2004) and this may also be the case here. A more definitive answer would require the analysis of BupsΦ1 at the genetic level, with comparison to bacteriophage genes of known function present in other bacteriophage.

A second bacteriophage extracted from another isolate of high virulence (Φ7, extracted from #7) was used for infection studies. The only experiment in which Φ7 was successfully inserted (Assay G.2), did not show alterations in virulence. This was not examined further as confirmation of insertion with this bacteriophage was unsuccessful in most cases.

Plaque forming assays were carried out on NCTC 13178, NAFC and selected isolates of NAFC+BupsΦ1 to glean some indication of the level of stability of the prophage. These assays indicated that the bacteriophage extracted from NCTC 13178 was

capable of stable lysogeny. Addition of mitomycin C did not cause enough lysis to be visible as plaques irrespective of the bacteria to agar ratio. As group two and three assays did in some cases produce plaques, it is reasonable to assume that plaques were present but difficult to see due to low bacterial growth in group one. As such, groups two and three only will be discussed. The first observation of note was that the addition of mitomycin C prior to plating did not appear to cause any changes in plaque production in agar.

Of most interest in this assay was the production of plaques on NAFC and not on NAFC+BupsΦ1 in group two. As NAFC is an isogenic mutant of NCTC 13178, it would be expected that if no plaques were seen in NCTC 13178, then no plaques would be seen in NAFC. The presence of plaques in NAFC could indicate several things. Firstly, the bacteriophage in NAFC mutated such that it was less stable in a lysogenic lifecycle than it was in NCTC 13178. Secondly, the plaques were due to a second bacteriophage which infected NAFC or finally, a second bacteriophage already existed which the bacteriophage in NCTC 13178 holds in a stable lysogenic state and a mutation in BupsΦ1 in NAFC altered this. During bacteriophage and DNA extraction and digestion of NAFC, no bands indicative of a second bacteriophage were found.

Evidence to support the first theory can be seen in test two (Table 5.2) with the loss of plaque production in NAFC+BupsΦ1 which had been reinfected with bacteriophage extracted from NCTC 13178 (assay A.1). If reinsertion of a fully functional bacteriophage occurs, that bacteriophage may produce enough of the relative gene products to retain lysogenic stability for both bacteriophage. The single plaque produced by NAFC+BupsΦ1 from assay B.1 in group two indicates that reinfection with bacteriophage may not always be sufficient to block plaque production.

In contrast, in group three (Table 5.2), NAFC+BupsΦ1 from assay A.1 was found to produce plaques. The only differences between group two and three assays were the

broth to agar ratio and the age of the isolate from which the broth was made. The broth in group three was inoculated from an isolate plated for 24 hours longer than the isolate used for the group two broth. The extra day on Ashdown agar may have provided an extra stress triggering lysis in this case. This extra day of stress did not cause lysis in NCTC 13178, nor in NAFC+BupsΦ1 from assay B.1. This could be explained by differences in insertion into the chromosome. As noted previously, some bacteriophage are capable of insertion into multiple sites in the bacterial chromosome, with bacteriophage P2 able to integrate into at least ten sites on *E. coli* (Barreiro and Haggard-Ljungquist, 1992) and bacteriophage Mu having a very large number of insertion sites (Hsu and Davidson, 1972; Manna *et al.*, 2004). If the insertion site in the current study differs from one experiment to the other and from the original insertion point, this could result in phenotypic changes as have been noted with bacteriophage Mu (Taylor, 1963; Hsu and Davidson, 1972). In addition the bacterial response to stress may also differ. It has been shown that bacteriophage insertion can inactivate chromosomal genes, such as is the case with the inactivation of β-lysin in *S. aureus* (Coleman *et al.*, 1989) and this may cause variation in the bacterial response to environmental cues.

The pattern of presence or absence of plaques may hint at possible interactions at the molecular level. However, to definitively determine whether insertion sites are playing a role in either virulence, as suggested above, or playing a role in stability of the offspring, the site of insertion would need to be determined. Prior to this, the bacteriophage would need to be sequenced to select DNA sequences to use for identification of the insertion region.

In conclusion, it is likely that BupsΦ1 from NCTC13178 successfully inserts into NAFC as there are permanent colonial morphology changes. It is possible that BupsΦ1 carries virulence genes as one experiment showed an upregulation of virulence. However, as this was not repeatable, and no increase in virulence of other isolates occurred, it is more likely that BupsΦ1 causes phenotypic changes in the bacterial host due to insertion site. This phenotypical change may in some cases

cause virulence changes, both increasing and decreasing. Variable lysis of cells may also play a role in virulence, with slower growth of bacteria a side effect of less stable bacteriophage.

Plaque forming assays (Table 5.2) indicated that while the bacteriophage in NAFC was less stable than that of NCTC13178, addition of BupsΦ1 from NCTC 13178 improved stability. It is possible that a gene such as *cII*, which at high levels represses the lytic cycle, has been damaged in the NAFC bacteriophage. Addition of functioning *cII* produced by BupsΦ1 would repress lysis of both bacteriophage. Alternatively, the *cI* gene encoding the repressor, CI, may be damaged in such a way that lysis is not as strongly repressed. In both scenarios, lysis of cells containing damaged as well as functional bacteriophage would occur more quickly than lysis of cells containing only functional bacteriophage, as was seen in test three.

CHAPTER 6

SEQUENCING OF BACTERIOPHAGE BupsΦ1

6.1 Introduction

The ability to identify prophage sequences in bacterial genomes is helpful in understanding the variation of individuals in a species and can aid in analysis of the evolution of both bacteriophage and their hosts (Casjens, 2003; Benson *et al.*, 2004). It can also aid in understanding bacteriophage-host interactions and potentially allow us to improve our strategies for curing diseases (Uetz *et al.*, 2004). These interactions can include imparting new genetic attributes to the host, including virulence factors. As discussed earlier, there are many examples of bacteriophage encoding toxins and other factors which may give the bacterial host an environmental advantage and may increase virulence (Beutin *et al.*, 1993; Rubin *et al.*, 1998; Bensing *et al.*, 2001; Boyd and Brussow, 2002).

Bacteriophage proteins such as endolysin, which can be used to break down bacterial cell walls (Boizet *et al.*, 1990), are being examined for possible therapeutic use. Identifying the genes producing these proteins and being able to recombinantly produce these proteins has potential benefits for health in a post antibiotic world. Already, loads of group B streptococci in the oropharynx and vagina of mice have been lowered using cloned lysin (Cheng *et al.*, 2005b). Similarly, lysin produced from the bacteriophage of *Bacillus anthracis* has been shown to have specific bacteriolytic activity against *B. anthacis*, both *in vitro* and *in vivo* (Schuch *et al.*, 2002).

Genetic modifications made to bacteriophage are being used to target particular treatments. Lysis deficient bacteriophage have been used experimentally to treat both murine peritonitis (Matsuda *et al.*, 2005) and *Pseudomonas aeruginosa* infections (Hagens *et al.*, 2004) with minimal production of endotoxin, hence minimising inflammation and increasing survival rates.

Understanding of bacteriophage sequence has also been of value in manipulation of the bacteriophage for use as a research tool. Bacteriophage λ is the most well known bacteriophage which has been used as a tool in molecular work (Sambrook and Russell, 2001). Knowing the genetic structure of this bacteriophage and being able to alter it to suit the purpose of the researcher has allowed for great flexibility. Typically, bacteriophage vectors are capable of insertions of DNA much larger than can be carried by plasmids. Whole genomes can be inserted (genomic DNA libraries) and expressed (expression libraries). More recently, genetically modified bacteriophage expressing antibodies on their surface have been used to block the effect of cocaine by acting as receptors for the cocaine (Carrera *et al.*, 2004; Nutt and Lingford-Hughes, 2004). Bacteriophage are of particular interest in this regard as they can cross the blood brain barrier and mop up cocaine in areas of the body where traditional medications have difficulty reaching. In a similar fashion, bacteriophage displaying antibodies to amyloid plaques present in the brains of Alzheimer's disease patients, have been examined in mice as a probe for the purpose of early identification of the disease (Frenkel and Solomon, 2002).

Analysis of bacteriophage genomic information has been undertaken using different tools, depending on the desired outcome. DeShazer (2004) used subtractive hybridization to identify a prophage in one strain of *B. pseudomallei* that was not present in another strain and from this information, a bacteriophage was extracted and the DNA analysed. Anne *et al.* (1984; 1985) extracted bacteriophage and purified the DNA, prior to carrying out restriction digest mapping and sequencing of bacteriophage VWB from *Streptomyces,* for the purpose of developing a cloning vehicle from the bacteriophage. This second method of extracting and purifying bacteriophage DNA is commonly carried out when the presence of an inducable bacteriophage has previously been confirmed. Typically, once DNA has been extracted, multiple restriction endonucleases are tested to select those which cut the bacteriophage into fragments of manageable size. As well as being used to calculate expected genomic size, these can be cloned into plasmids and sequenced using universal primers. Sequences can then be joined up, either by primer walking,

shotgun PCR matching of sequences, BLAST searching to determine expected position and orientation relative to other sequences prior to selective PCR, or by multiple restriction endonuclease sequence generation and overlapping of the sequences generated. Often a combination of these methods is applied (Anne *et al.*, 1985; Oakey *et al.*, 2002; Braid *et al.*, 2004; Summer *et al.*, 2004; Kwan *et al.*, 2005; Van Dessel *et al.*, 2005).

Several potential problems with restriction endonuclease cutting exist. Methylation of DNA, where present, can protect the DNA from being cut with restriction endonucleases, resulting in slow or partial digestion (Anne *et al.*, 1984). In addition, selection of restriction enzymes for digestion of DNA, such that resultant band sizes are appropriate for cloning work, is dependant on the bacteriophage. This may require many restriction endonucleases being trialed before one is selected (Anne *et al.*, 1984; Oakey *et al.*, 2002).

Initial analysis of DNA is commonly carried out using an ORF finder and a BLASTp search to compare potential genes (Oakey *et al.*, 2002; Braid *et al.*, 2004; Summer *et al.*, 2004; Kwan *et al.*, 2005; Van Dessel *et al.*, 2005). Nucleic acid sequences exhibit much larger sequence heterogeneity compared to their encoded protein sequences due to redundancy of the genetic code, so use of amino acid coding when comparing and aligning sequences is desirable to remove a level of this heterogeneity (Stocsits *et al.*, 2005). Heterogeneity can also exist when a protein can perform the same task and have the same chemo-physical properties but otherwise have very different amino acid coding. Several mono-ADP-ribosyltransferases (ADP-RTs) that act as toxins, such as exotoxin A of *P. aeruginosa*, diphtheria toxin of *Corynebacterium diphtheriae* and cholera toxin of *V. cholerae* have conserved amino acid coding with regions of dissimilar amino acids between conserved residues (Wozniak *et al.*, 1988). These long regions of dissimilar amino acids resist any automated alignment. This means identification of genes across species can be problematic without already having some idea of what gene you are looking at. It is relevant to note that ADP-RT toxins have been found encoded on bacteriophage in

some organisms (Greenfield *et al.*, 1983; Popoff *et al.*, 1990; Waldor and Mekalanos, 1996) and that *B. pseudomallei* has been recorded as producing an ADP-RT toxin (Mohamed *et al.*, 1989b).

In order to understand how BupsΦ1 interacts with its host, it is useful to determine the possible proteins encoded on the bacteriophage. From this, identification of any toxin genes or transcriptional regulator genes can be done. Sequencing of the BupsΦ1 genome is an effective way of determining potential proteins and identifying their relevance with respect to virulence. Since previous work (Chapter 4) has identified BupsΦ1 as belonging to the family *Myoviridae*, structural genes and others similar to those already identified in other *Myoviridae* bacteriophage can be anticipated. Similar gene order is likely, with phage genomes typically showing gene clustering according to function (Casjens, 2003). Remaining open reading frames can be examined with respect to similarities to reported toxin genes or transcriptional regulator genes. Chemo-physical comparison can be undertaken where unknown genes are of a correct size for feasible toxin genes of known amino acid diversity.

6.2 Materials and Methods

6.2.1 Bacteriophage isolation and DNA extraction

6.2.1.1 Production of DNA for restriction digestion

Bacteriophage BupsΦ1 was isolated and amplified as per Section 4.2.2 using *B. pseudomallei* #4 as the propagating strain. One hundred millilitre volumes of broth grown to an $O.D._{600nm}$ of 0.1 (or about $1x10^8$cfu/ml) was used and $1x10^9$pfu of bacteriophage was added. Filtered, amplified solutions of BupsΦ1 were concentrated by ultracentrifugation as per Section 4.2.2.2. The pellet was resuspended in 500ul of SM buffer modified to contain 0.4M potassium chloride at a pH of 6.3 (Appendix 1). Any DNA not bound inside capsid was removed by binding to activated extraction matrix (Nucleobond AX kit; Machery-Nagel, Düren, Germany) as follows. Five hundred µl of buffer L1 (Nucleobond AX kit; Machery-Nagel, Düren, Germany) was added to 0.05g of extraction matrix. This was gently inverted for five minutes at room temperature using an end over end mixer (Laboratory suspension mixer,

Clements Pty. Ltd, NSW, Australia) to activate the matrix. The solution was then centrifuged at 5000*g* for five minutes at room temperature. The supernatant was discarded and the bacteriophage sample added. This was mixed by gentle inversion for 20 minutes and then centrifuged at 5000*g* for five minutes at room temperature. The supernatant was removed without any carry over of matrix and kept for DNA extraction.

DNA extraction was performed as per Section 4.2.3 without DNase 1 addition. Sheared DNA was removed by processing the sample through a gel clean up or PCR clean up kit (Perfectprep® Gel Cleanup kit, Eppendorf, Hamburg, Germany; QIAquick® PCR Purification kit, Qiagen, Germany; HiYield Gel/PCR DNA Fragments Extraction Kit, Real Biotech Corp., Taiwan, ROC.; Wizard® SV Gel and PCR Clean-Up System, Promega, Madison, USA) to a final volume of 300µl.

6.2.2 Sequence determination

6.2.2.1 Restriction digestion and cloning

Restriction endonucleases (Promega, Madison, USA) were selected to include four, six and eight base cutters and to be compatible with pBK-CMV (Stratagene, USA). Each restriction endonuclease was tested for suitability using one µg of BupsΦ1 DNA according to suppliers' instructions. Restriction enzymes were determined to be suitable if bands of digestion fragments of between 600 and 3000 base pairs were produced on a gel.

Digestion fragments of between 600 and 3000 base pairs were excised from the gel using clean scalpel blades. DNA fragments were removed from the agarose slices using DNA gel extraction kits (Perfectprep® Gel Cleanup kit; HiYield Gel/PCR DNA Fragments Extraction Kit; Wizard® SV Gel and PCR Clean-Up System) according to the kit's instructions. DNA was eluted into deionised water (ddH_2O), $O.D._{260/280nm}$ was determined as per Section 3.2.5 and the sample was stored at -20°C until required.

The same restriction endonuclease was used to cut pBKCMV according to suppliers' instructions and the ends dephosphorylated (Section 3.2.3.1). Fragments were ligated into the dephosphorylated plasmid and transformed into chemically competent *E. coli* cells as per Sections 3.2.3.2 and 3.2.3.3.

Presumptive transformants were selected on the basis of blue-white screening, with white colonies indicative of an insert present in the plasmid. Up to three presumptive transformants from each ligation were further cultured overnight for plasmid extractions (Section 3.2.4). The presence of an insert was confirmed by taking 500ng of plasmid, linearising it by making a single cut using a restriction enzyme present in the plasmid multi-cloning site (other than the restriction endonuclease used to digest the bacteriophage DNA) and examining the size of the linear product using gel electrophoresis (Section 3.2.7). Alternately, the presence of an insert was confirmed by carrying out a PCR using universal primers present on the plasmid, a general PCR mix (Table 3.3) and the PCR program designed for sequencing (Section 3.2.9). The size of the product was then examined using gel electrophoresis.

Plasmids with confirmed inserts were either sent to Macrogen (Korea) for sequencing or processed at James Cook University. At James Cook University, plasmids were sequenced using a sequencing PCR program (Section 3.2.9) with plasmid sequencing primers (Table 6.1), and Amersham DYEnamic ET Teminator Cycle Sequencing Kit (Amersham Biosciences, NSW, Australia). All oligonucleotides were synthesised by Sigma Aldrich, Australia. Reactions were analysed at the Genetic Analysis Facility (GAF) at James Cook University, Townsville, QLD, Australia, using a Beckman CEQ 2000 DNA Analysis System or an Amersham Biosciences MegaBASE 1000 .

In all cases, data was examined using Sequencher™ software (Gene Codes Corporation, Ann Arbor, MI, USA) and any vector and primer sequence was removed. Fragment sequences were aligned and overlapped where possible using Sequencher™, to form a number of contiguous data sets (contigs). Where sequenced

inserts were too long to produce a single contig with one sequencing event, primer walking was carried out to join sequences together.

Primer walking involved the creation of oligonucelotide primers internally on the insert, directed towards the unknown central region of the insert. This was done using DNA sequence already determined as a template and carrying out sequencing using the new primers. Oligonucleotide primers were designed using Oligo v 6.65 (Section 3.2.8). Further rounds of primer walking were carried out until a single contig of each insert sequenced was achieved.

Table 6.1 Primers used for sequencing inserts in plasmids. Primer sequence is shown in 5' to 3' direction. Some primer names are the same between brands but have different sequence structure

Plasmid name and Brand	Primer name	Primer sequence
pBK-CMV (Stratagene; USA)	BK reverse M13 (-20)	ACAGGAAACAGCTATGACCTTG GTAAAACGACGGCCAGT
pBK-CMV	T7 T3	GTAATACGACTCACTATAGGGC ATTAACCCTCACTAAAGGGA
pGEM-T (Promega, USA)	T7 SP6	TAATACGACTCACTATAGGG TATTTAGGTGACACTATAG
pGEM-T (Promega, USA)	M13F M13R	GTTTTCCCAGTCACGAC CAGGAAACAGCTATGAC
T&A cloning Vector (Real Biotech Corp., Taiwan, ROC)	M13F M13R	TTTCCCAGTCACGAC TCACACAGGAAACAGCTATGAC

6.2.2.2 Primer directed approach

Oligonucleotide primers were designed to amplify out from both ends of all contigs. Primers were designed with similar melting temperatures using the primer design component of Oligo v6.65. All combinations of primers, with the exception of those produced at either end of a single contig, were tested using conventional PCR conditions (Section 3.2.4.3). PCR mixes were developed in plate format using the CAS1200 liquid handling robot (Corbett Robotics, Qld, Australia) and products were

visualised using gel electrophoresis. PCRs which resulted in either single strong bands of <1500 bases (or a strong band of <1500 bases as well as one or more faint bands) were run on gel electrophoresis and the strong band cut out and cleaned using a gel clean up kit (Perfectprep® Gel Cleanup kit, Eppendorf, Hamburg, Germany; HiYield Gel/PCR DNA Fragments Extraction Kit, Real Biotech Corp., Taiwan, ROC.; Wizard® SV Gel and PCR Clean-Up System, Promega, Madison, USA). PCR products were ligated into pGEM-T (Promega) or yT&A (Yeastern Biotech, Taiwan) using the supplier's recommended protocol. Plasmids were transformed and confirmed for insert as noted earlier (Section 6.2.2.1). Plasmids with confirmed inserts were sequenced using plasmid universal primers (Section 6.2.2.1). Fragment sequences were aligned and overlapped where possible to existing contigs using Sequencher™. All sequence data was subjected to a BLASTx (translated sequence vs. protein database) search of the GenBank database.

To estimate position and orientation of large contigs and fragments along the genome, long range PCR was applied to combinations of primers using whole DNA from BupsΦ1 as a PCR template. Instead of standard *Taq* polymerase, platinum *pfx* (Invitrogen, Australia) was used according to supplier's instructions. Two reaction cycle conditions were examined, long product conditions, to amplify extremely long products and five kb conditions, to amplify products up to five kb. In both cases primers were designed to have a melting temperature of above 70°C and were RP1 purified.

1. Long product conditions

- Denaturation at 94°C for five minutes
- 15 cycles of;
    1. denaturing at 94°C for 30 seconds
    2. annealing and extension of primers at 72°C for 25 minutes
- 15 cycles of;
    1. denaturing at 94°C for 30 seconds
    2. annealing and extension of primers at 72°C for 25 minutes plus 15 seconds per cycle

- one cycle of;
    1. denaturing at 94°C for 30 seconds
    2. annealing and extension of primers at 72°C for 40 minutes
- storage at 4°C in the thermocycler until removed and stored at 4°C prior to further processing.

2. Five kb product conditions

- Denaturation at 94°C for five minutes
- 30 cycles of;
    1. denaturing at 94°C for 30 seconds
    2. annealing and extension of primers at 72°C for three minutes
- one cycle of;
    1. denaturing at 94°C for 30 seconds
    2. annealing and extension of primers at 72°C for ten minutes
- storage at 15°C in the thermocycler until removed and stored at 4°C prior to further processing

Products were visualised using 0.8% agarose gel electrophoresis. Products which exceeded the minimum expected size were cut out and either directly sequenced using the primers used in the PCR or cloned into a T-vector after A-tailing. Sequence fragments which matched the ends of the contigs from which the primers were designed were considered proof that the PCR product would join the two contigs tested, as well as confirmation of the orientation and size of gap between the two contigs.

A-tailing was carried out by making up the following reaction mix to 75µl in a thin walled 500µl PCR tube and heating the reaction to 72°C for ten minutes (Ray Layton, JCU, pers. com.). The reaction was then cooled to 15°C and cleaned by use of a PCR cleanup kit (HiYield Gel/PCR DNA Fragments Extraction Kit, Real Biotech Corp., Taiwan, ROC).

Reaction mix
PCR product (after gel extraction)
1 x reaction buffer (Promega)
50µM dATP (Promega)
2.5mM $MgCl_2$ (Promega)
0.05 units/µl *Taq* polymerase (Promega)

The linking of contigs using PCR continued until sufficiently large contigs for open reading frame (ORF) analysis were obtained. At least two replicates of each base were determined, through either overlapping of contigs or replicate testing of fragments. Where ambiguities occurred in duplicate sequences, at least one more reaction was undertaken to determine the most likely correct base.

6.2.3 Open reading frame (ORF) analysis

Nucleotide sequence data for BupsΦ1 was analysed for open reading frames (ORFs) and putative genes using NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) with bacterial coding selected in parallel with a BLASTx of GenBank. Results of the ORF Finder analysis were applied to a BLASTp (protein sequence vs. protein database search) of GenBank to assign a hypothetical role or function and thus form a putative gene map of the BupsΦ1 genome. This was compared to the results of the BLASTx to identify any likely start/stop codon errors. The resultant putative genes were isolated and applied to a final BLASTx to confirm putative identification, E value and % identity.

Where no match or a hypothetical match of unknown origin was detected, ORF nucleotide sequence was translated into amino acid sequence using BCM Search Launcher: Six Frame Translation of Sequence (http://searchlauncher.bcm.tmc.edu/seq-util/Options/sixframe.html). Where the product size of ORF was reasonably similar to the product of interest, the correct frame amino acid sequence was then examined for the existence of expected amino acid order or chemo-physical properties to match to either ADP-RT or endolysin.

Analysis of ORFs for presence of an endolysin was carried out using ClustalX (1.8) (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/) to align amino acid translations of ORF sequences and GeneDoc v.2.6.002 to visualise chemo-physical relatedness of aligned sequences. Multiple ORF alignments were carried out, the poorest matches removed and new alignments carried out until only the most likely matches remained. Alignments were also visualised using NJplot (http://pbil.univ-lyon1.fr/software/njplot.html) (Perriere and Gouy, 1996) with bootstrap values included (number of bootstrap trials = 1000).

Analysis of amino acid translations of ORFs for the presence of conserved amino acid residues indicative of ADP-RTs was carried out by examining all ORFs in ClustalX for expected amino acid residues and examination of the spacing between residues. Two amino acid patterns were used as a template to look for ADP-ribosylating activity in the ORFs. The first was histidine or arginine...105 to 130 amino acids...glutamic acid...one amino acid...glutamic acid (**H**/**R**...105 to 130aa...**E**...1aa...**E**), with a hydrophobic and or aromatic region 50-75aa downstream of the **H**/**R**. This was based on the identification of the conserved regions of group one ADP-RTs (Takada *et al.*, 1995). The second was the pattern **R**...50aa to 120aa...**E**...1aa...**E** with the conserved region serine-threonine-serine (**STS**) in the 50 to 120aa variable region. This was based on the conserved active sites of group four ADP-RTs reported by Barth *et al* (1998). Patterns were searched for using a string search in ClustalX by recording all amino acids fitting the necessary matches and comparing spacing between these with spacing of the exotoxins. All ORFs were examined as some possible ADP-ribosylating activity has previously been recorded in genes already putatively identified to have a function (Oakey *et al.*, 2005).

6.3 Results

6.3.1 Effectiveness of DNA preparation methods

6.3.1.1 DNA extraction and cleanup

All kits used for the cleanup of DNA from bacteriophage extraction were found to produce a similar quality and quantity of DNA. The yield of DNA after cleanup

where DNase 1 was not used was approximately one quarter the initial yield. Most of the DNA lost was due to removal of sheared, low molecular weight DNA. One hundred millilitres of bacteriophage broth typically produced a final concentration of about 100µg/ml DNA or a total of 30µg of DNA. Comparison on a gel showed little difference before and after cleanup in the intensity of the high molecular weight band. Where DNase 1 was used (Section 4.2.3), the final yield was considerably lowered, to about three µg of DNA from 100ml of bacteriophage broth. This was probably due to fragile neck structures allowing some entry of DNase 1 to the capsid. This amount was insufficient for restriction digestion for the purpose of sequencing of bands.

6.3.1.2 Restriction digestion

Eighteen restriction endonucleases, including two combinations of multiple restriction endonucleases were tested for suitability (Table 6.2). Of these, only three were determined to be suitable, *Eco*RV+*Sma*I, *Sal*I and *Pst*I. All four base cutting restriction endonucleases caused complete digestion of the bacteriophage DNA.

Table 6.2 Grouping of restriction endonucleases by cutting pattern. Non-cutters are those which produced only one high molecular weight band by gel electrophoresis. Poor cutters produced few or no bands within the desired size range of 600bp to 3000bp, but may produce several bands above this size range. Good cutters produced five or more bands within the required size range. Complete cutters caused complete digestion of DNA. Information in brackets beside the name of the restriction endonuclease describes whether it is an unusual (U), four (4), six (6), or eight (8) base cutter and whether it cuts a blunt end (B).

Non Cutters	Poor cutters	Good Cutters	Complete Cutters
*Bgl*II(6) *Dra*I(6,B) *Xba*I(6) *Eco*RI(6) *Spe*I(6)	*Eco*RV (6,B) *Sma*I(6,B) *Not*I(8) *Kpn*I(6) *Sfi*I(U) *Bam*HI(6) *Eco*RV(6,B)+*Sfi*I(U)	*Pst*I (6) *Sal*I(6) *Eco*RV(6,B)+*Sma*I(6,B)	*Sau*3A I(4) *Hae*III(4,B) *Hpa*II(4)

6.3.1.3 Gel extraction

The quantity of DNA in a band was calculated using a marker containing quantitative bands (GeneRuler™ 1Kb DNA Ladder, Fermentas, USA) and Genetools software (Gentools v.3.00.22, Synoptics, England). All gel extraction kits trialed were similar in efficiency with a final yield of about 50% of the DNA being extracted. Yield was increased slightly when elution buffer or water used for elution was warmed to 50°C prior to addition to column and was allowed to sit for one minute in the column prior to elution.

6.3.1.4 Ligation and cloning of extracted restriction fragments

Restriction endonucleases that produced blunt ends were examined first due to the flexibility of ligation into any blunt end cut plasmid. Only one restriction fragment of seven was successfully ligated into pBK-CMV using this method in spite of addition of ten units of T4 DNA ligase to the ligation mix (Section 3.2.3). Fifteen fragments produced using *Sal*I (Figure 6.1) were successfully ligated into pBK-CMV which had been cut with *Sal*I and dephosphorylated. Sequencing of multiple clones identified some bands fragments as consisting of multiple fragments of the bacteriophage genome.

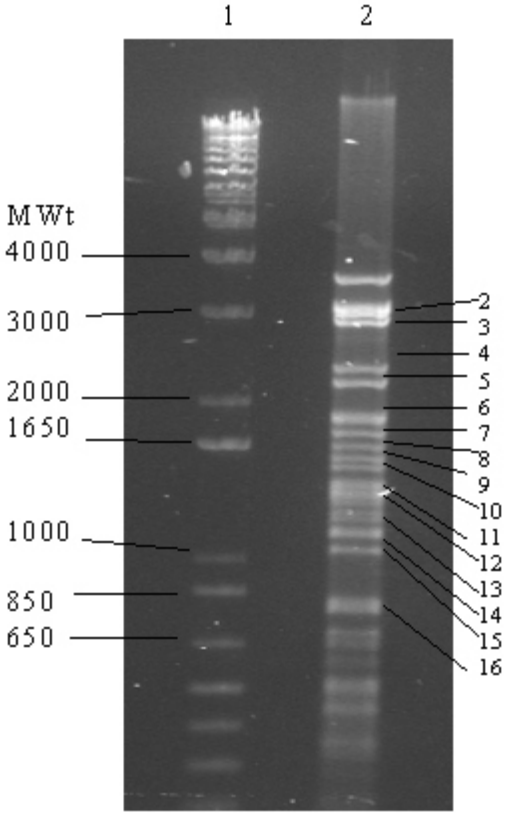


Figure 6.1 Gel of *Sal*I digest of bacteriophage BupsΦ1. Lane 1; 1kb plus DNA ladder (Invitrogen, USA), Lane 2; *Sal*I digested BupsΦ1. Bands which were cut out and sequenced are numbered 2-16.

6.3.1.5 Primer directed approaches

Shotgun sequencing of BupsΦ1, using combinations of primers designed from either end of all contigs, mainly resulted in the extension of single contigs or the production of new contigs. These new contigs did not match previously designed contigs due to some level of non-specific binding. Long amplification primers were designed from the two longest contigs produced from sequencing cloned restriction. These were trialed in all combinations using both the long amplification PCR and the five kb PCR. In each combination, multiple bands were produced, some of these bands were similar in both PCR protocols, while others were unique. Only one combination of primers produced a bright band which was large enough to be the specific product, with the size of this bright band differing in each PCR. In each case the band was excised and sequenced. Both bands were found to be specific to and extend one

contig, but not match to the other. The five kb PCR produced a longer product (3000 bp) and was used for further attempts to combine contigs.

Few of the attempts to join contigs resulted in expected contigs being joined. However, the length of the products meant even non-specific products produced useful lengths of DNA and several contigs were combined after non-specific products were produced. All primers designed to extend contigs 7 and P1 produced only non-specific binding, unrelated to these contigs. As specific products could not be guaranteed, combination of all contigs into a single genome was delayed and there was not sufficient time or funding to complete the genome during the period of the experiment. As such, the ORFs of all contigs are separately listed in Table 6.3. Contig nomenclature varies as labelling was altered for different batches of restriction fragment and primer directed sequencing, as well as alteration when contigs were combined.

6.3.2 Analysis of ORFs

Those contigs with putative ORFs relating to DNA replication are recorded first (Table 6.3), followed by those encoding putative structural genes. The expected order of the structural genes of a myovirus are as follows;

Terminase-portal-protease-scaffold-majorheadshell(coat) protein - head/tail joining proteins-tail shaft protein-tapemeasure protein-tail tip/baseplate proteins-tailfibre (Casjens, 2003).

Analysis of ORFs often resulted in strong matches to putative genes on contigs of *B. vietnamiensis* bacterial genome, which is still incomplete at the time of writing. Closer examination of these ORFs showed that some of them were matches to putative bacteriophage genes in the *B. vietnamiensis* genome and all contigs carrying

these ORFs also carried putative bacteriophage genes based on BLASTx matches. This indicated that these strong matches were likely to be to a prophage in the *B. vietnamiensis* bacterial genome. Where matches for different bacteriophage contigs had hits to different areas of the same *B. vietnamiensis* contig, these contigs were recorded in Table 6.3 in the same order. In cases where multiple ORFs on the bacteriophage contig matched to ORFs on a *B. vietnamiensis* contig, it was observed that the direction and order of the genes also matched. Furthermore, two pairs of contigs (A and B) and (D and U) that matched to a single *B. vietnamenis* contig were later combined via primer directed sequencing, confirming direction and order of contigs (AB), (DU). A visual representation of other not yet combined matches to *B. vietnamiensis* contigs is displayed in Figure 6.2.

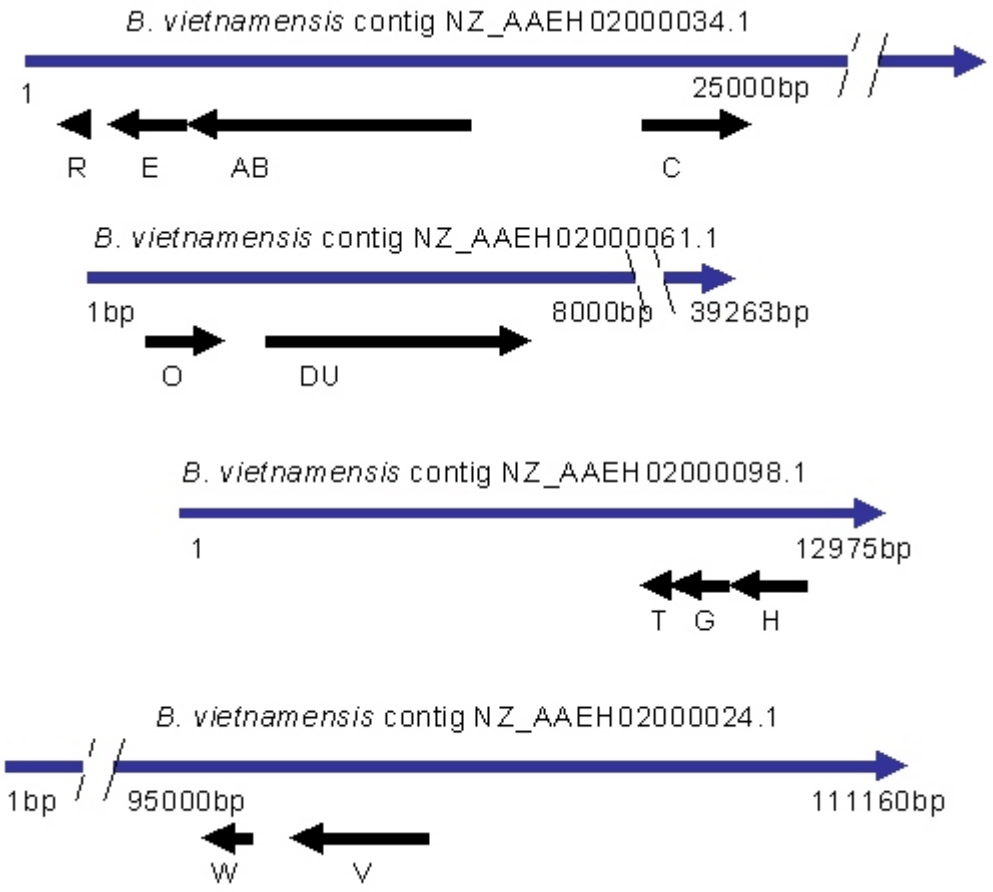
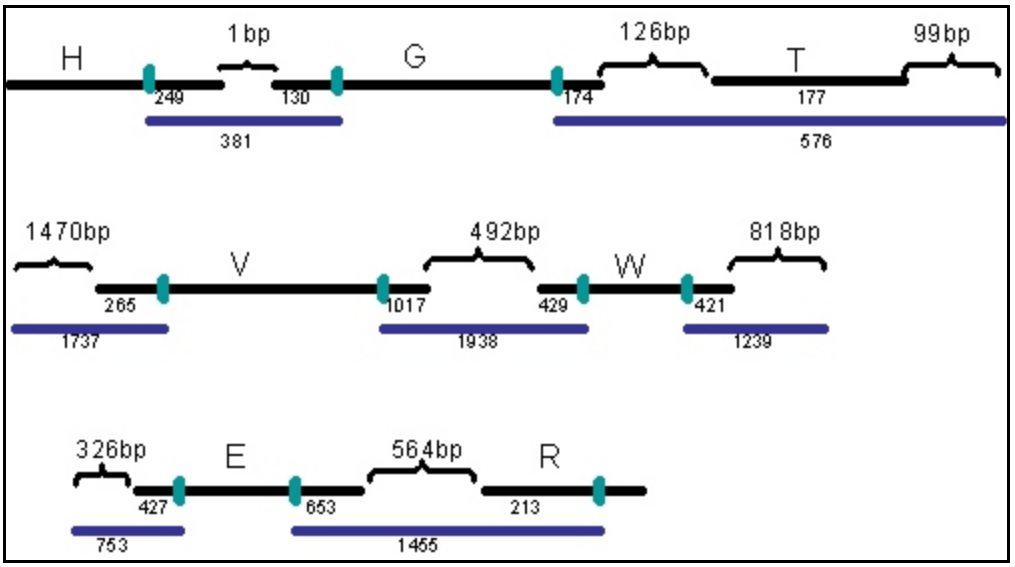


Figure 6.2 Representation of probable contig direction and order based on ORF matches to *B. vietnamiensis* contigs. *B. vietnamiensis* contigs are represented by blue arrows with the contig name above, BupsΦ1 contigs are represented by black arrows with the contig name below. Images are not to scale.

Several contigs were difficult to interpret. Contigs 7 and P1 may have been bacterial contamination as they could not be extended using subsequent bacteriophage DNA extracts. Contig I carried a kanamycin resistance gene identical to that found in many plasmids and may have been PCR contamination from a plasmid not used in this study. Other contigs, primarily consisting of one or two partial ORFs matching strongly to bacterial genomes and not matching to *B. vietnamiensis* contigs, were considered to be suspect non-phage contigs until such time as they can confirmed by completion of sequencing of the genome. Contigs 7, P1, I, 11RFsmF, S, L, DX, 2FR and Dr-15b3F are suspected non-phage. They are placed at the end of Table 6.3 below a dividing line (shaded row). Their ORF analysis has been included. Where included in data analysis they are described as suspect contigs or suspect ORFs.

The partial ORFs at the end of contigs were included (Table 6.3) where a strong match to part of gene was found and the match was logical, i.e. where the match was to the appropriate end of the gene. Where two contigs have partial ORFs at their ends matching either end of a gene, these contigs are placed next to each other in the appropriate order. A visual representation of these partial matches is displayed in Figure 6.3. All other contigs follow in no particular order. Suspect ORFs as noted in the preceding paragraph are not included in Figure 6.3.

A



B

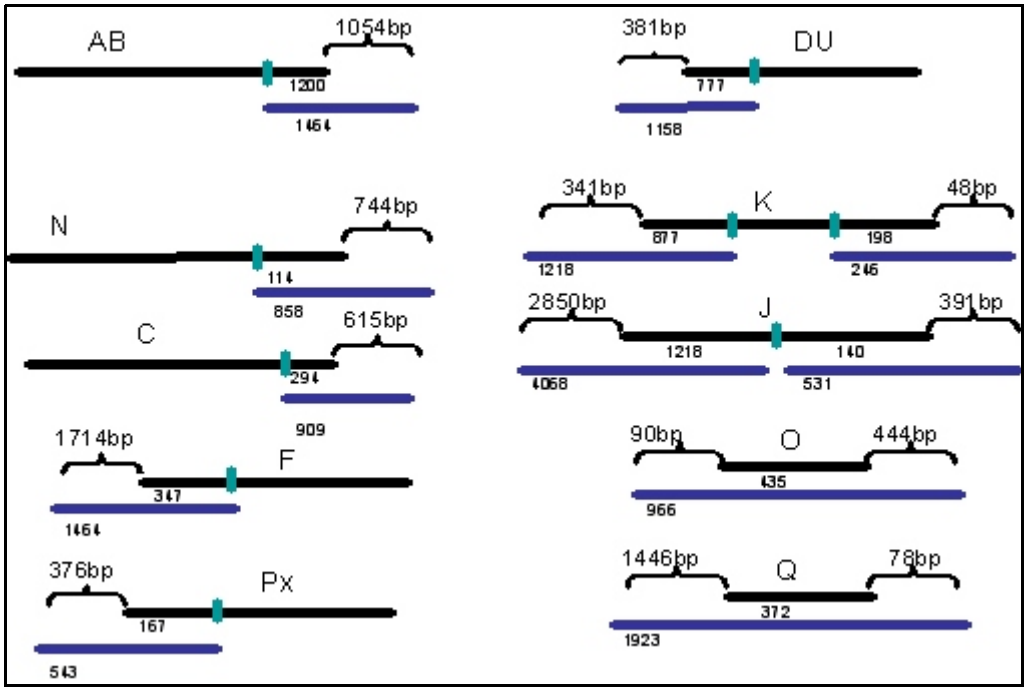


Figure 6.3 Representation of probable unsequenced bases based on ORF matches to genes recorded in GenBank using BLASTx. Genbank genes are represented by blue lines with base pair length below, BupsΦ1 contigs are represented by black lines with name above. The position of relevant ORFs are represented by vertical aqua lines and base pair length of ORFs is noted under contig. Calculated range of unsequenced genome in base pairs is noted above bracket. Figure 6.3A consists of all contigs which have matches to genes which also match another contig. Figure 6.3B displays all contigs with single gene matches at their ends. E value and % identity are recorded in Table 6.3. Images not to scale.

Table 6.3 Putative open reading frames (ORFs) identified in contigs of the BupsΦ1 genome. E value and % identity information gathered from BLASTx analysis after position was confirmed using ORF finder, BLASTx analysis of entire contigs and examination of the contig sequence in Sequencher™. ORFs with '*' have nearest protein sequence identified by BLASTp as BLASTx did not produce a hit. Nucleotide length (bp: base pairs) includes stop codon. Amino acid length (aa) does not. Where partial ORFs at ends of two contigs match to either end of the same gene, these were placed sequentially and ORFs are highlighted in grey. Contig order is based on relative order of matches to *B. vietnamiensis* contigs and expected order of structural ORFs. All contigs where this is not relevant are placed in no particular order below these. Suspect contigs are listed at the end of the table, below the greyed out row.

Contig name and size	***Open reading frames***	***Position***	***Length bp***	***aa***	***Nearest protein sequences on GenBank (identified by BLASTx)***	***E-value (Identity %)***
C 5846bp	C1	complement 4-306	303	100	ZP_00425300.1: hypothetical protein Bcep1808DRAFT_2309 *Burkholderia vietnamiensis* G4. 99 aa	1e-43 83%
	C2	complement 382-591	210	69	EAO09455.1: Transaldolase subfamily *Nocardioides* sp. JS614. 446 aa	0.13 36%
	C3	complement 683-1024	342	113	no match	
	C4	1008-1538	531	176	NP_945073.1: gp42 Bacteriophage phi1026b. 142 aa	3e-50 74%
	C5	1535-1792	258	85	NP_945072.1: gp41 Bacteriophage phi1026b. 85 aa	4e-41 96%
	C6	1785-2027	243	80	NP_945071.1: gp40 Bacteriophage phi1026b. 80 aa	6e-28 73%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	C7	2017-2598	582	193	ZP_00494949.1: COG0568: DNA-directed RNA polymerase, sigma subunit (sigma70/sigma32) *Burkholderia pseudomallei* Pasteur. 70 aa	6e-19 94%
	C8	2595-2897	303	100	ZP_00425300.1: hypothetical protein Bcep1808DRAFT_2309 *Burkholderia vietnamiensis* G4. 99 aa	1e-43 83%
	C9	3017-3229	213	70	ZP_00427889.1: hypothetical protein Bcep1808DRAFT_7740 *Burkholderia vietnamiensis* G4. 110 aa	1e-12 57%
	C10	3190-3414	225	74	YP_164266.1 hypothetical protein F116p02 *Pseudomonas aeruginosa* phage F116. 160 aa	6e-5 41%
	C11	3436-3744	309	102	no match	
	C12	3737-3862	126	41	ZP_00425302.1: hypothetical protein Bcep1808DRAFT_2311 *Burkholderia vietnamiensis* G4. 87 aa	5e-14 94%
	C13	3859-4098	240	79	NP_569248.1: hypothetical protein HCM1.24c *Salmonella enterica subsp. enterica* 106 aa	0.015 32%
	C14	complement 4139-4375	237	78	ZP_00461461.1: hypothetical protein Bcen2424DRAFT_4619 *Burkholderia cenocepacia* HI2424. 72 aa	2e-4 43%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	C15	complement 4425-4634	210	69	AAR23187.1: gp36 Bacteriophage phi1026b. 69 aa	2e-15 52%
	C16	complement 4673-4861	189	62	no match	
	C17	complement 4858-4944	87	29	no match	
	C18	4943-5179	237	78	no match	
	C19	5176-5556	381	126	YP_164270.1: hypothetical protein F116p06 *Pseudomonas aeruginosa* phage F116. 213 aa	5e-14 37%
	C20	5553-post	294+	98+	ZP_00279444.1: COG4422: Bacteriophage protein gp37 *Burkholderia fungorum* LB400. 303 aa	3e-34 67%
AB 13865bp	AB1a	57-122	66	21	no match	
	AB1b	129-248	120	39	no match	
	AB1ca	339-657	319	106	no match	

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	AB1c	complement 707-805	99	32	no match	
	AB1d	complement 847-906	60	19	no match	
	AB1e	complement 891-972	81	26	no match	
	AB2	970-1269	300	99	YP_355356.1: gp21 *Burkholderia cepacia* phage Bcep176. 110aa	2e-7 42%
	AB3	complement 1196-1552	357	118	NP_536409.1: putative repressor protein Bacteriophage phiE125. 130aa	3e-6 47%
	AB4	1628-1939	312	103	YP_006380.1: Cro *Salmonella typhimurium* bacteriophage ST104. 71aa	0.12 54%
	AB5	2056-2430	375	124	ZP_00425226.1: hypothetical protein Bcep1808DRAFT_2235 *Burkholderia vietnamiensis* G4. 123aa	3e-58 87%
	AB6	2427-2651	225	74	ZP_00425225.1: hypothetical protein Bcep1808DRAFT_2234 *Burkholderia vietnamiensis* G4. 74aa	6e-26 88%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	AB6a	2605-2850	246	81	no match	
	AB7*	2869-3159	291	96	ZP_00493053.1: hypothetical protein BpseP_01003527 *Burkholderia pseudomallei* Pasteur. 68aa	3e-4 40%
	AB8	3135-4025	891	296	ZP_00425396.1: hypothetical protein Bcep1808DRAFT_2819 *Burkholderia vietnamiensis* G4. 221aa	3e-8 36%
	AB9	4067-5158	1092	363	NP_536413.1: putative DNA cytosine methylase Bacteriophage phiE125. 378aa	0.0 95%
	AB10	5266-6117	852	283	NP_536414.1: putative PAPS reductase/sulfotransferase Bacteriophage phiE125 gp57. 332 aa	9e-86 56%
	AB11	6114-6722	609	202	ZP_00623574.1: hypothetical protein NhamDRAFT_3513 *Nitrobacter hamburgensis* X14. 318 aa	3e-25 51%
	AB12	complement 6703-7962	1260	419	Domain I: ABA50693.1: hypothetical protein BURPS1710b_1799 *Burkholderia pseudomallei.* 621aa (Fits in 106-550aa -complementary)	1.6 28%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
					Domain II: CAE49823.1: Putative DNA-binding protein *Corynebacterium diphtheriae*. 238aa (595-783bp of AB12)	2.6 38%
	AB12a	7961-8047	87	28	no match	
	AB13	8044-8514	471	156	ZP_00425222: Endodeoxyribonuclease RusA *Burkholderia vietnamiensis* G4. 158 aa	2e-57 81%
	AB14	8559-8852	294	97	ZP_00425221: hypothetical protein Bcep1808DRAFT_2230 *Burkholderia vietnamiensis* G4. 109 aa	4e-17 60%
	AB15	8849-9559	711	236	YP_355337: gp2 *Burkholderia cepacia* phage Bcep176. 196 aa	4e-79 74%
	AB16	complement 9546-9821	276	91	CAC 90909: hypothetical phage protein *Yersinia pestis* CO92. 91 aa	1e-7 34%
	AB16a	complement 9827-9910	84	27	no match	
	AB17	9909-10574	666	221	CAC90914: hypothetical phage protein *Yersinia pestis* CO92. 211 aa (Or possible transposase)	3e-59 55%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	AB18	10525-11073	549	182	CAC90915: hypothetical phage protein *Yersinia pestis* CO92. 149 aa	5e-26 50%
	AB19	11075-12670	1596	531	AAT37736: gp33 TerL *Burkholderia cenocepacia* phage BcepB1A. 532 aa	5e-153 51%
	AB20	12666-post	1200+	405+	AAP49949.1: putative portal protein *Enterobacteria* phage T1. 427 aa	2e-12 23%
E	E1	5-427	423	140	ZP_00425217.1: Phage putative head morphogenesis protein, SPP1 gp7 *Burkholderia vietnamiensis* G4. 251 aa	5e-67 97%
2306bp	E2	429-722	294	97	ZP_00350984.1: COG0243 Anaerobic dehydrogenases, typically selenocysteine-containing *Ralstonia eutropha* JMP134. 120 aa	2e-14 41%
	Elow	643-951	309	103	no match	
	E3	955-1476	522	173	NP_284620.1: hypothetical protein NMA1913 *Neisseria meningitidis* Z2491. 197 aa	1e-14 31%
	E4	complement 1423-1553	130	43	no match	

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	E5	1622-post	685+ 228+	ZP_00425216.1: conserved hypothetical protein *Burkholderia vietnamiensis* G4. 485 aa fits in gene at 1-228aa (10-684bp of E5)	1e-100 89%
R 227bp	R1	pre-213	213+ 71+	ZP_00425216.1: conserved hypothetical protein *Burkholderia vietnamiensis* G4. 485 aa fits in gene at 416-485aa (end of gene)	2e-30 94%
Q 376bp	Q1	5-post	372+ 123+	ZP_00710967.1: COG5525: Bacteriophage tail assembly protein *Escherichia coli* B171. 641 aa	2e-54 99%
O 435bp	O1	all 435 bases	435+ 145+	NP_852772.1: hypothetical protein Aaphi23p45 Bacteriophage Aaphi23. 284 aa	1e-6 31%
DU 4870bp	DU1	pre-777	777+ 258+	ZP_00500901.1: COG3464: Transposase and inactivated derivatives *Burkholderia pseudomallei* S13. 386 aa	2e-129 100%
	DU2	862-1215	354 117	ZP_00168251.1: COG3573: Predicted oxidoreductase *Ralstonia eutropha* JMP134. 116 aa	1e-33 76%
	DU3	1212-2390	1182 393	ZP_00724345.1: COG3299: Uncharacterized homolog of phage Mu protein gp47 *Escherichia coli* F11. 399 aa	7e-96 48%
	DU4	2393-3055	663 220	ZP_00426840.1: putative bacteriophage protein *Burkholderia vietnamiensis* G4. 220 aa	7e-108 87%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	DU5	3116-4537	1422	473	Domain I: NP_805636.1: putative bacteriophage tail protein *Salmonella enterica* subsp. *enterica* serovar Typhi Ty2. 503 aa	2e-11 (53%)
					Domain III: AAS47891.1: gp52 *Burkholderia cenocepacia* phage BcepMu. 785 aa	2e-79 48%
	DU6	4547-4729	183	60	ZP_00827773.1: hypothetical protein YfreA_01003992 *Yersinia frederiksenii* ATCC 33641. 60 aa	4e-6 37%
	DU7	4770-post	100+	33+	no match	
V 2618bp	V1	pre-265	265+	88+	YP_296580.1: hypothetical protein Reut_A2374 *Ralstonia eutropha* JMP134. 579 aa	5e-19 65%
	V2	243-842	600	199	no match	
	V2.5	972-1412	441	146	YP_296576.1: hypothetical protein Reut_A2370 *Ralstonia eutropha* JMP134. 164 aa	4e-54 68%
	V3	1418-1612	195	64	ZP_00490831.1: hypothetical protein Bpse6_01000298 *Burkholderia pseudomallei* 668. 61 aa	2e-6 38%

Table 6.3 continued

Contig name and size	***Open reading frames***	***Position***	***Length bp aa***	***Nearest protein sequences on GenBank (identified by BLASTx)***	***E-value (Identity %)***
	V4	1599-post	1017+ 338+	ZP_00494945.1: COG0270: Site-specific DNA methylase *Burkholderia pseudomallei* Pasteur. 646 aa (fits in gene at 19-313aa)	0.0 97%
W 862bp	W1	pre-429	429+ 143+	ZP_00494945.1: COG0270: Site-specific DNA methylase *Burkholderia pseudomallei* Pasteur. 646 aa (fits in gene at 458-570aa)	2e-74 96%
	W2*	442-post	421+ 140+	YP_111061.1: hypothetical bacteriophage protein *Burkholderia pseudomallei* K96243. 413 aa	0.13 33%
H 1502bp	H1	pre-721	721+ 240+	ZP_00986362.1: COG4834: Uncharacterized protein conserved in bacteria *Burkholderia dolosa* AUO158. 363 aa	1e-118 92%
	H2	723-1190	468 155	ZP_00427731.1: hypothetical protein Bcep1808DRAFT_0012 *Burkholderia vietnamiensis* G4. 155 aa	3e-32 83%
	H3	1254-post	249+ 82+	ZP_00427730.1: putative bacteriophage protein *Burkholderia vietnamiensis* G4. 127 aa	8e-19 92%
G 1568bp	G1	pre-130	130+ 43+	ZP_00427730.1: putative bacteriophage protein *Burkholderia vietnamiensis* G4. 127 aa	4e-16 97%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp*	*aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	G2	159-638	480	159	ZP_00168238.1: COG0034: Glutamine phosphoribosylpyrophosphate amidotransferase *Ralstonia eutropha* JMP134. 159 aa	2e-53 89%
	G3	670-933	264	87	no match	
	G4	1018-1389	372	123	ZP_00427746.1: putative bacteriophage protein *Burkholderia vietnamiensis* G4. 123 aa	2e-63 94%
	G5	1394-post	174+	57+	ZP_00427745.1 putative bacteriophage protein *Burkholderia vietnamiensis* G4. 192 aa (fits in at 1-58aa of gene)	7e-11 93%
T 177bp	T	All 177bp - part of gene	177+		ZP_00427745.1: putative bacteriophage protein *Burkholderia vietnamiensis* G4. 192 aa (fits in at 100aa-152aa)	1e-25 96%
K 1231bp	K1	pre-877	877+	292+	YP_331533.1: gp30 *Burkholderia pseudomallei* 1710b. 406 aa	5e-153 98%
	K1a	899-973	75	25	no match	
	K2	970-1074	105	34	AAR31204.1: unknown *Burkholderia pseudomallei*. 34 aa	5e-12 100%
	K3	1034-post	198+	65+	ZP_00496820.1: COG3514: Uncharacterized protein conserved in bacteria *Burkholderia pseudomallei* S13. 90 aa	2e-20 98%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
F 2071bp	F1	complement pre-328	328+ 109+	no match	
	F2	complement 354-542	189 62	no match	
	F3	483-776	294 97	YP_355373.1: gp38 *Burkholderia cepacia* phage Bcep176. 91 aa	5e-6 37%
	F4	complement 957-1154	198 65	no match	
	F5	1378-1554	177 58	YP_355376.1: gp41 *Burkholderia cepacia* phage Bcep176. 71 aa	0.02 44%
	F6	1551-1700	150 49	ZP_00490837.1: hypothetical protein Bpse6_01000304 *Burkholderia pseudomallei* 668. 51 aa	7e-6 60%
	F7	1697-1840	144 47	no match	
	F8	complement 1815-post	257+ 85+	no match	

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
J 1366bp	J1	pre-1223	1218+ 384+	ZP_00501774.1: COG0046: Phosphoribosylformylglycinamidine (FGAM) synthase, synthetase domain *Burkholderia pseudomallei* S13. 1356 aa	0.0 100%
	J2	1228-post	139+ 46+	ZP_00437794.1: COG3153: Predicted acetyltransferase *Burkholderia mallei* 10399. 177 aa	0.81 100%
Px 434bp	Px1	pre-167	167+ 54+	AAN16057.1: ORF182 *Pseudomonas stutzeri.* 181 aa	1.1 38%
	Px2	complement 224-376	153 50	AAR23202.1: gp51 Bacteriophage phi1026b. 50 aa	4e-17 84%
N 493bp	N1	pre-70	70+ 23+	no match	
	N2	69-371	303 100	CAH39632.1: transposase *Burkholderia pseudomallei* K96243. 100 aa	8e-38 100%
	N3	380-post	114+ 38+	CAH39631.1: putative IS element protein *Burkholderia pseudomallei* K96243. 286 aa	5e-14 97%
				SUSPECT CONTIGS BELOW THIS LINE	

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
7 1772bp	7-1	complement pre-544	544+ 180+	ZP_00890775.1| COG0075: Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase *Burkholderia pseudomallei* 1106a. 404 aa (fits 1-171aa of gene)	2e-82 100%
	7-2	527-880	354 117	AAU49590.1: hypothetical protein BMA2048 *Burkholderia mallei* ATCC 23344. 117 aa	5e-32 98%
	7-3-a	complement 909-968	60 29	no match	
	7-3	979-1236	258 85	XP_521221.1: PREDICTED: similar to eukaryotic translation initiation factor 4B *Pan troglodytes*. 354 aa (91-234bp of 7-3 fits 154-200aa of gene)	2.4 39%
	7-4	complement 1311-1421	111 36	no match	
	7-5	complement 1442-post	331+ 109+	YP_109329.1: putative LysR-family transcriptional regulator *Burkholderia pseudomallei* K96243 aka COG0583. 326 aa (fits 1-110aa of gene)	2e-54 100%
P1 1201bp	P1-1	pre-544	544+ 181+	ABA47516.1: hypothetical protein BURPS1710b_2381 *Burkholderia pseudomallei* 1710b. 630 aa	3e-88 100%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	P1-2	complement 498-848	351 116	YP_333775.1: 50S ribosomal protein L31 *Burkholderia pseudomallei* 1710b. Also COG0254. 116 aa	2e-43 100%
	P1-3	complement 976-post	225+ 74+	ZP_00478729.1: COG3228: Uncharacterized protein conserved in bacteria *Burkholderia pseudomallei* 1710a. 281 aa	4e-37 100%
I 1390bp	I1	complement 49-282	177 58	no match	
	I2	345-1154	810 269	ABB04083.1: KanR Shuttle vector pMQ61. 269 aa	4e-159 100%
11RF-smF 1053bp	RF1	pre-664	664+ 221+	EAO07842.1: hypothetical protein NocaDRAFT_3156 *Nocardioides* sp. JS614. 310 aa	1.6 25%
	RF2	658-768	111 36	no match	
	RF3	805-post	249+ 83+	YP_432251.1: predicted phosphoesterases, related to the Icc protein *Hahella chejuensis* KCTC 2396. 252 aa	7.1 49%
S 219bp	S1	whole contig	219+ 73+	NP_921778.1: putative gypsy-type *retrotransposon Oryza sativa* japonica cultivar-group. 1995 aa 13-219bp matches 780-848 aa	8e-10 46%

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
L 816bp	L1	pre-135	135+ 44+	BAA82345.1: ORF1 TT virus. 53 aa	1.5 43%
	L2	200-307	108 35	No match	
	L3	338-post	479+ 159+	ZP_00168213.2: COG1674: DNA segregation ATPase FtsK/SpoIIIE and related proteins *Ralstonia eutropha* JMP134. 292 aa (fits 8-166aa of gene)	5e-47 54%
DX 968bp	DX1	pre-409	409+ 136+	ZP_00168213.2: COG1674: DNA segregation ATPase FtsK/SpoIIIE and related proteins *Ralstonia eutropha* JMP134. 292 aa (fits 162-285aa of gene)	3e-38 59%
	DX2	406-post	563+ 188+	ZP_00168212.2: COG1196: Chromosome segregation ATPases *Ralstonia eutropha* JMP134. 579 aa (fits 1-131aa of gene)	1e-49 57%
2FR 1689bp	2FR1	complement pre-1071	1071+ 357+	ZP_00497647.1: COG0719: ABC-type transport system involved in Fe-S cluster assembly, permease component *Burkholderia pseudomallei* S13. 488 aa 1071-1 matches to 1-358 of gene	0.0 98%
	2FR middle	1156-1587	432 143	no match	

Table 6.3 continued

Contig name and size	*Open reading frames*	*Position*	*Length bp aa*	*Nearest protein sequences on GenBank (identified by BLASTx)*	*E-value (Identity %)*
	2FR2	1550-post	141+ 46+	ZP_00475477.1: COG3213: Uncharacterized protein involved in response to NO *Burkholderia pseudomallei* 1710a. 380 aa (Fits last 46aa of gene)	5e-21 100%
Dr-15b3F 1279bp	DR-1	all 1279 bases	1279+ 426+	ZP_00479442.1: COG3210: Large exoproteins involved in heme utilization or adhesion *Burkholderia pseudomallei* 1710a. 3049 aa	1e-129 99%

6.3.3. Examination of contigs

As the genome of BupsΦ1 was not completely sequenced, multiple contigs remain. In some cases, a partial ORF at the end of one contig will match to one end of a gene and a partial ORF at the end of a second contig will match to the other end of the same gene (Figure 6.3A, Table 6.3). Extrapolation of this data can give a good indication of the likely number of base pairs of DNA that would exist between contigs where contig ends match the same gene. If the unsequenced regions between these contigs is identical to the reported gene to which the partial ORFs match, the combined unsequenced regions between contigs with matching ends is 1183 bp.

An extrapolation can also be carried out for those contigs with partial ORFs at their ends where there is not a second contig that matches to the same gene. The number of extra bases extrapolated from this data, not including suspect contigs, is 13657 bp. The total number of sequenced bases recorded in Table 6.3, not including suspect contigs, is 40913 bp. Addition of the sequenced and calculated bases results in 55753 bp. The previously determined size of the genome (Section 4.3.7) is 55142 bp., 611 bases less than the size calculated by sequencing and gene extrapolation.

6.3.3.1 Structural genes

Putative structural genes were found on contigs AB, E, Q and DU. These include AB19: TerL (terminase large subunit), AB20: portal protein, E1: partial ORF - head morphogenesis, Q: partial ORF - tail assembly, DU5: tail fibre (Figure 6.4). The contig order from AB to DU is AB, E, R, Q, O, DU (Table 6.3). Contigs R and O consist of partial ORFs which match to hypothetical undescribed genes. These were placed in this order as AB-R has strong matches to a single *B. vietnamiensis* contig (Figure 6.2) as do contigs O-DU. Logically, the unsequenced regions and some or all of the hypothetical unidentified ORFs between AB19 and DU5 encode those structural genes not identified above.

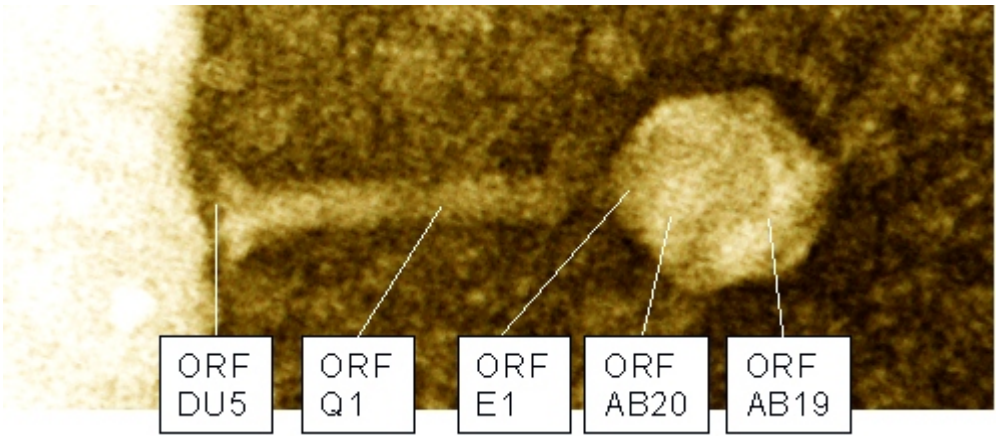


Figure 6.4 Image of BupsΦ1 for purpose of matching putative structural genes to relevant structures. Putative genes responsible for structural development are shown. Image shows from left to right; a probable portion of bacterial cell wall, tail fibres, uncontracted sheath, neck and capsid. DNA remains in capsid as shown by uncontracted state of sheath and pale colouration of capsid. After DNA release, the capsid would be observed as a dark hexagon. Putative genes are as follows. Tail fibres; ORF DU5, tail assembly; ORF Q1, head morphogenesis; ORF E1, portal protein; ORF AB20, terminase large subunit; ORF AB19

6.3.3.1.1 Putative tail fibres

The putative tail fibre genes have partial matches to different genes (Figure 6.5). There appear to be three domains in ORF DU5, domain I (1-219 bp, 73 aa) correlates to the first 73 aa of a 503 aa putative tail fibre in the *S. enterica* genome (53% match) and a 348 aa hypothetical protein in *R. eutropha* (71% match) respectively. Domain II (220-267 bp, 16 aa) consists of a unique region, while domain III (268-1419 bp, 384 aa) correlates to the right hand portion of gp52 of *B. cenocepacia* phage BcepMu (785 aa, 47% match), a P2 gpH tail fiber protein. Domain III has a slightly shorter correlation (268-1320 bp, 350 aa) to the right hand portion of COG3210 (471 aa, 34% match), a large exoprotein involved with either adhesion or haeme utilization in *B. pseudomallei.*

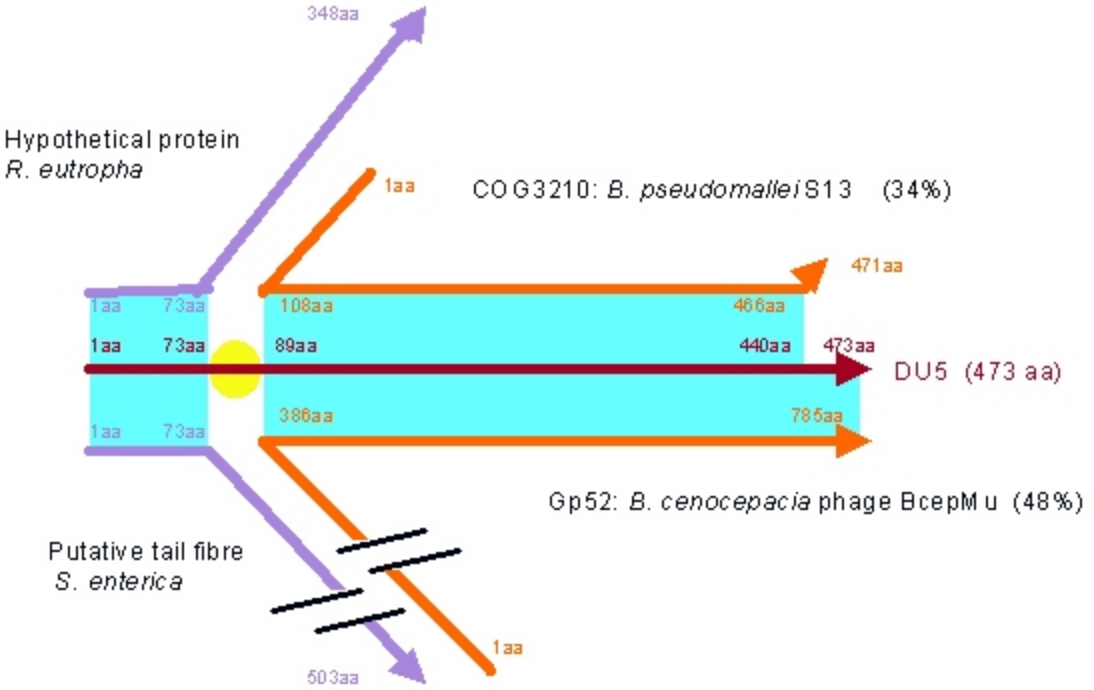


Figure 6.5 Alignment of putative tail fibre gene ORF DU5 (in red) to genes found in GenBank. Genes matching to Domain I of DU5 are shown in lilac, genes matching to Domain III of DU5 are shown in orange. The short unique region of DU5 is circled in yellow. Relevant amino acid residues, including position 1 and final amino acid as well as position of start and end of match, are shown in the same colour as the gene. Domains I and III are shaded. Accession numbers for matching genes are; Hyp. *R. eutropha*: YP_296618.1, Putative tail fibre *S. enterica*: NP_805636.1, COG3210 *B. pseudomallei* S13: ZP_00497081.1, gp52 BcepMu: AAS47891.1

6.3.3.2 DNA replication and lysogeny

Putative genes involved in DNA replication and lysogeny were found on contig AB and contig C. Contig C carried two ORFs (C1 and C8) that were inverted repeats of each other, indicative of transposon activity. Examination of the regions around these ORFs determined that the inverted region extended another 116 bp from the three prime end of C1 and the five prime end of C8. The other end of the inverted region could not be identified directly as C1 lies at the five prime end of the contig. A possible 11 bp short inverted repeat region was identified being just interior to the

right hand end of the first region (C1+) at bp 409-419. An inverted repeat of this was found at the left end of the second region (C8-) at bp 2482-2492 of the contig and a direct copy was found at what was assumed to be the right of region C8 at bp 5315-5325. This would make the inverted repeat regions 2822 bp long, not including short repeat regions. Additional PCR and sequencing within this region confirmed that the information was correct and not a sequencing error. Four other ORFs or partial ORFs (AB17, DU1, N1, N2) were identified as transposases or insertion elements by BLASTx homology. None of the complete ORFs (AB17, N1), had the short repeat or inverted repeat regions expected of transposons.

Contig C also carries an RNA polymerase sigma subunit (ORF C7), probably responsible for transcription of late genes. The second inverted repeat region includes part of this ORF, with part of the match to the RNA polymerase gene being in this region. ORF AB9 may be a DNA cytosine methylase, for methylating bacteriophage DNA and protecting it from host defences. Contig N carries a transposase (ORF N1) as well as a partial ORF which may be an integrase. Contig AB also carries ORFs matching to RusA, a Holliday junction resolvase (AB13); a repressor protein (AB3); and a weak match to Cro (AB4). These exist for the purpose of resolving Holliday junctions for insertion of DNA into the host genome, repressing the lytic lifecycle and triggering lysis respectively.

6.3.3.3 Virulence genes

Several ORFs putatively encoded for possible virulence components. Of these, ORF E2 has been putatively identified as an anaerobic dehydrogenase. Anaerobic dehydrogenases are enzymes necessary for various bacterial functions in anaerobic environment such as many *in vivo* environments. ORF DU2 was predicted to be an oxidoreductase, an enzyme that protects bacteria from the oxidative damage caused by eukaryotic host defences against intracellular bacteria.

One of the virulence components of interest is that of an ADP-ribosylating exotoxin. No ORF was identified by homology searching to match any recorded exotoxin. All ORFs, including suspect ORFs, were examined for conserved amino acid residue

patterns typical of ADP-RTs. The patterns described a glutamic acid (**E**)-any amino acid (**X**)- glutamic acid (**E**) conserved region, at least 50aa (group two to four) or 105aa (group one) downstream of a highly conserved arginine (**R**) or histidine (**H**) residue, with the second glutamic acid vital to ADP-ribosylation activity (Barth *et al.*, 1998). The first glutamic acid is not always conserved. It is not present in exotoxin A of *P. aeruginosa* (a group one ADP-RT)(Gray *et al.*, 1984), so may not be essential to function in all cases. Twenty-one ORFs contained at least one iteration of the **EXE** region at least 50aa in from the start of the gene. This number was cut to eight when selection was based on a 105aa residue distance. None of the 21 ORFs had residues downstream from the **EXE** region (**EXE**-tyrosine(**Y**)-isoleucine(**I**) or **EXEXXX**-tryptophan(**W**)) that are found in some ADP-ribosylating toxins. Of the eight likely ORFs, seven had at least one **R** or **H** 105-130 residues upstream of the **EXE** region, as are found in group one exotoxins. ADP-RTs have also been described as having a hydrophobic or aromatic region 50 to 75aa downstream from the **R** residue (Takada *et al.*, 1995). The seven candidate ORFs were examined to determine if any had this region present based on the example in Takada *et al*, (1995). Two ORFs were found to have 25aa regions of 60% or more hydrophobic and or aromatic residues. ORF AB10 has the pattern $\mathbf{H}_{(8)}$-49aa- 60% hydrophibic region$_{(58\text{-}83)}$-42aa-$\mathbf{E}_{(126)}$-1aa-$\mathbf{E}_{(128)}$, ORF K1 has the pattern $\mathbf{R}_{(6)}$-74aa-72% hydrophobic region$_{(80\text{-}105)}$-16aa-$\mathbf{E}_{(122)}$-1aa-$\mathbf{E}_{(122)}$ (Figure 6.6).

Group two to four exotoxins also contain a highly conserved serine-threonine-serine (**STS**) region between the **R** and **EXE** region and group two toxins have an addition conserved **H** residue between the **R** and **STS** motifs. All ORFs were examined for the presence of the **STS** motif. Six ORFs had **STS** motifs (AB12, C3, C13, D5, F1, F4), however C3, C13 and F4 were eliminated as the **STS** motif lay too close to the right hand end of the sequence. D5 was also eliminated as there was no **E** downstream of the **STS** motif. Only AB12 also had an **EXE** region but this was incorrectly positioned relative to the **STS** region. Where present, the **EXE** pattern has always been reported as having only one amino acid between the two glutamic acid residues. AB12 and F1 were examined more closely with respect to common group two and four residues and the following was determined;

ORF F1 had an amino acid residue pattern of **R**$_{(17)}$-29aa-**STS**$_{(47-49)}$-30aa-**E**$_{(80)}$ (Figure 6.6). This is one of the smallest recognised toxin patterns, but within the limits of those recorded. The best associations of AB12 with ADP-ribosylating toxins were with the residue patterns **R**$_{(122)}$ -24aa- **H**$_{(147)}$-22aa- **STS**$_{(170-172)}$ - 48aa-**E**$_{(221)}$ and **R**$_{(5)}$-43aa-**H**$_{(49)}$-11aa-**lysine(K)TS**$_{(59-61)}$-51aa-**E**$_{(113)}$-1aa-**E**$_{(115)}$. As lysine is very different in coding, structure and chemo-physical properties from serine, this second amino-acid pattern is unlikely to be an active toxin. There was another **E** five residues downstream of the reported **E**$_{(221)}$ but as the single amino acid gap between these two **E** residues seems to be conserved in ADP-ribosylating toxins, this glutamic acid residue has been discounted as being involved.

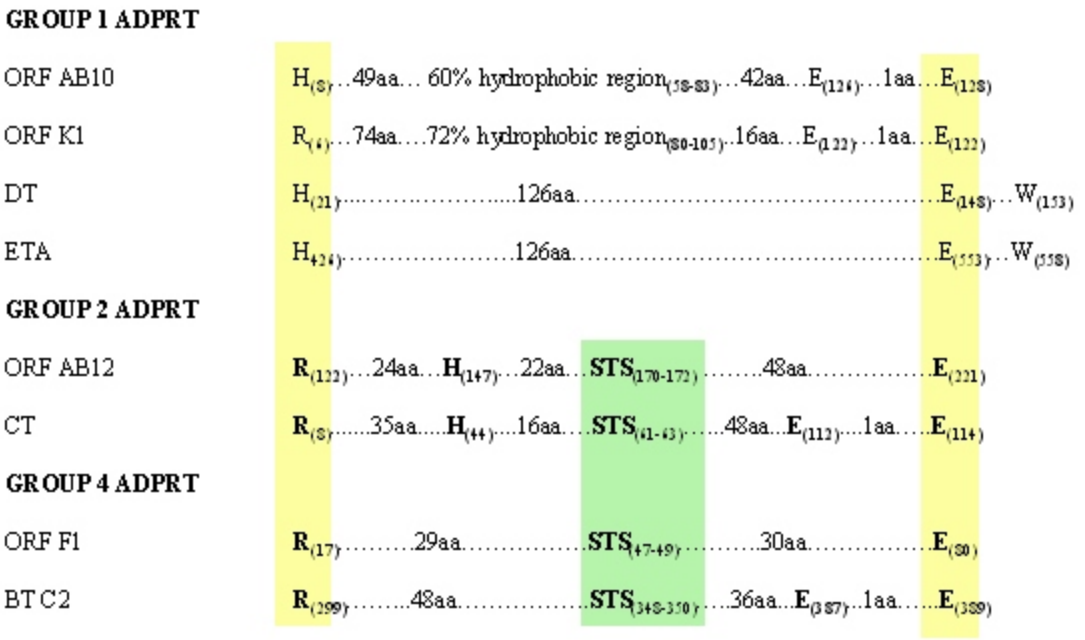


Figure 6.6 Comparison of selected open reading frames with several recorded ADP ribosyltransferases. DT: diphtheria toxin (Greenfield *et al.*, 1983), ETA: exotoxin A of *P. aeruginosa* (Wozniak *et al.*, 1988), CT: cholera toxin (Yamamoto *et al.*, 1984), BT C2: Botulinum toxin C2 (Fujii *et al.*, 1996; Barth *et al.*, 1998). Yellow shading is conserved arginine/histidine R/H) and glutamic acid (E) respectively. Green shading is conserved STS motif. Subscripts in brackets represent position of amino acid. ...'X'aa... represents number of amino acids between conserved regions

6.3.3.4 Lysis genes

No homology to an endolysin gene was determined via a search of GenBank. An examination of endolysin genes of *Burkholderia* bacteriophage already recorded in GenBank (Table 6.4) using ClustalX alignment and chemo-physical matching in GeneDoc v.2.6.002 (www.psc.edu/biomed/genedoc) indicated that the genes shared strong chemo-physical alignment, with genes ranging in size from 127 aa to 255 aa. Initially, all unidentified ORFs (including suspect ORFs) were compared to these using ClustalX and GeneDoc. The poorest matches, as ascertained by eye, were removed until two possible ORFs (DU3:393 aa and DU4:220 aa) were selected as having high chemo-physical matches to the endolysin gene (gp28) of *Burkholderia cenocepacia* phage Bcep1. Of these two ORFs, DU3 is much larger than any of the lysin genes examined and has multiple unique regions in the match. DU4 is closer to an expected lysin size. The recorded gene with greatest chemo-physical similarity to DU4 is gp28 of *Burkholderia cenocepacia* phage Bcep1 (NP_944336.1). This ORF is in a region expected to contain structural (tail) genes.

Examination of ORF AB15, AB17 and AB18 was also carried out using this method. Lysin genes have been recorded as being present upstream of TerL (AB19) and these ORFs were of sufficient size, did not have a function putatively ascribed, and were just upstream of AB19. Of these, ORF AB15 matched most closely to the recorded endolysin genes. Subjectively, the best match was to gp28 of *Burkholderia cenocepacia* phage Bcep1 (NP_944336.1). This match appeared to be better than that of DU4 to the same endolysin sequence. A phylogenetic tree with bootstrap scores was developed to provide a less subjective analysis. This tree examined the recorded genes and all the above-mentioned ORFs. No bootstrap score above 70% was found between the unknown ORFs and the recorded genes. The best score was 33.1% between ORF AB15 and a group of endolysin sequences (YP_109165.1, YP_024909.1, YP_355393.1, NP_536381.1, NP_945054.1, YP_333098.1).

An alignment of all ORFs in the genome with the known endolysin sequences and production of a phylogenetic tree displayed no clear relationship between the ORFs

and the endolysin sequences. The best bootstrap score was 16.8% between ORF 7-2 and NP_944336.1 (gp28 *Burkholderia cenocepacia* phage Bcep1) which is unlikely to be significant.

All unknown ORFs and those ORFs upstream of predicted possible lysin genes were examined in the same way with comparison to holin genes of bacteriophage of related hosts (Table 6.4). Examination of alignments and chemophysical properties in GeneDoc did not readily identify a potential holin gene. Bootstrap scores of the developed phylogenetic tree indicated that no ORFs examined had a probable (above 70%) relationship with any recorded holin gene. The best association was a 22% score between ORF AB1d and YP_024910.1 (gp04 S *Burkholderia cenocepacia* phage BcepB1A). Comparison of holin sequences with AB14 generated no bootstrap score.

Table 6.4 List of genes published on GenBank for comparison to ORFs of interest

Gene	geneID	protein ID	name	size (amino acid length)
Holin	gi|53718780	YP_107766.1	putative phage-related protein [*Burkholderia pseudomallei* K96243]	97
	gi|77864684	YP_355394.1	gp59 [*Burkholderia cepacia* phage Bcep176]	70
	gi|48697552	YP_024910.1	gp04 S [*Burkholderia cenocepacia* phage BcepB1A]	112
	gi|38707972	NP_945113.1	gp82 [Bacteriophage phi1026b]	118
	gi|38707913	NP_945053.1	gp23 [Bacteriophage phi1026b]	70
	gi|23752340	NP_705655.1	gp30 [*Burkholderia cepacia* phage Bcep781]	115
	gi|38638638	NP_944339.1	gp31 [*Burkholderia cenocepacia* phage Bcep1]	115
	gi|34610147	NP_918972.1	putative holin protein domain [*Burkholderia cepacia* phage BcepNazgul]	127
	gi|17975232	NP_536427.1	putative class I holin [Bacteriophage phiE125]	118
	gi|17975185	NP_536380.1	putative class II holin [Bacteriophage phiE125]	70
endolysin	gi|53720179	YP_109165.1	hypothetical protein BPSL2568 *Burkholderia pseudomallei* K96243	171
	gi|76811859	YP_333098.1	gp24 *Burkholderia pseudomallei* 1710b	163
	gi|77864683	YP_355393.1	gp58 *Burkholderia cepacia* phage Bcep176	165
	gi|48697551	YP_024909.1	gp03 R *Burkholderia cenocepacia* phage BcepB1A	165
	gi|38707914	NP_945054.1	gp24 Bacteriophage phi1026b	163
	gi|23752338	NP_705653.1	gp27 *Burkholderia cepacia* phage Bcep781	255
	gi|38638635	NP_944336.1	gp28 *Burkholderia cenocepacia* phage Bcep1	255
	gi|17975186	NP_536381.1	putative lysozyme Bacteriophage phiE125	127
	gi|34610146	NP_918971.1	putative endolysin *Burkholderia cepacia* phage BcepNazgul	134

6.4 Discussion

6.4.1 Restriction digestion, sequencing and combination of contigs

Of the 18 restriction endonucleases and combinations of endonucleases that were tested (Table 6.2), only three were considered suitable for further work. Of these one was a combination of two poor cutters. When these endonucleases were combined, cutting sites on the bacteriophage were shown to be of sufficient number and spacing to produce bands of the required size on a gel. However, ligation of these bands proved problematic. These restriction endonucleases produced blunt ends which were less efficiently ligated than cohesive or "sticky" ends (Pheiffer and Zimmerman, 1983). Two sticky ended restriction endonucleases were also considered good cutters. Of these *Sal*I was succesfully used to produce bands for ligation. Attempts to ligate bands digested with *Pst*I were not successful with one exception. As the quality and concentration of the bands was similar to the *Sal*I digestion, this is assumed to be due to the quality of the dephosphorylation step (Section 3.2.3.1) of the *Pst*I cut plasmid (pBK-CMV).

All four base restriction endonucleases were found to completely digest the BupsΦ1 DNA. Typically, the shorter the cut sequence, the more times that sequence will appear in a genome. This means shorter base cutters will more completely cut the genome as was shown here. Four base cutters were included in the testing of restriction endonucleases as previous genomes have required four base cutters for any digestion to be seen (Oakey and Owens, 2000). In this previous case, four base cutters were required to overcome the resistance to restriction digestion caused by methylation of the bacteriophage DNA (Oakey *et al.*, 2002), which is a common strategy of bacteriophage to overcome host defence mechanisms (Paul *et al.*, 2002). This result indicated the BupsΦ1 genome may not be effectively protected by methylation in spite of the presence of putative methylases (ORF AB9 and ORFs V4-W1).

Several primer directed approaches were used in an attempt to combine contigs. No technique was found to be highly successful in combining contigs, although expansion from one end of the contig sometimes resulted. The high level of non-

specificity in primer binding was probably due to the lack of sequence data on BupsΦ1 available when primers were designed. In support of this, as the amount of available sequence increased, fewer primer options were presented when primer design was carried out, indicating the first round of primer design had produced many non-specific primers.

Two of the original contigs created by restriction endonuclease digestion (P1 and 7), were not extended or combined. Primers designed for this purpose were invariably non-specifically bound when tested with primers designed from other contigs. One conclusion that may be made is that these contigs are the product of host bacterial contamination. In each digestion, an entire DNA extract was consumed and further primer directed amplification of sequence was carried out with different extracts. Should any host bacterial contamination exist, it could be of different portions of the bacterial genome in each case, thus providing the primers no DNA to specifically bind to. These contigs had no matches to identified bacteriophage genes; instead only having matches to genes identified as bacterial. This does not show conclusively that these contigs are the product of host bacterial contamination, but until the BupsΦ1 genome is completed, the ORFs on these contigs should not be considered to belong to BupsΦ1.

Contig I carried a kanamycin resistance gene which was a perfect match to many plasmid *kanR* genes. *B. pseudomallei* has kanamycin resistance; the plasmid used to generate this contig did not carry such a gene and the other ORF identified on this contig did not match to any other plasmid gene (or any gene). The perfect match to a plasmid kanamycin gene rather than a bacterial kanamycin gene however, indicates plasmid contamination of the PCR may be responsible for the product generated.

The DNA extraction method used for production of DNA for restriction digestion did not use DNase 1 to remove extraneous bacterial DNA present in the bacteriophage media. Using DNase 1 was found to lower the yield of bacteriophage DNA to three μg of DNA from 100ml of broth. This yield was insufficient for restriction digestion for the purposes of excision of digested bands. It also resulted

in difficulty in visualising bands. As the extraction protocol took four days, repeated extractions to increase yield were impractical. Instead, the extraneous DNA was removed using activated matrix from the Nucleobond AX kit previously described. This method produced 30µg of DNA from 100ml of broth (1/3 to 1/4 of the yield extracted when no method of DNA removal was carried out prior to extraction). Extraneous DNA was adsorbed to a silica matrix prior to separation from the bacteriophage and digestion of the bacteriophage capsid. For bacterial DNA to remain it must be bound to the structure of the bacteriophage in some way, possibly entangled with it. In using this method, the bacterial DNA removal step was increased from ten minutes to 20 minutes to minimise this possibility, but it cannot be completely discounted. Most of the primer directed approaches also used DNA extracted in this fashion as a template.

Given the possibility of contamination suspected from the analysis of contigs 7, P1 and I, other contigs, primarily consisting of one or two partial ORFs matching strongly to bacterial genomes, were labelled as suspect until such time as they can be confirmed by complete sequencing of the genome. Given the possibility of bacterial contamination, those PCR products carrying unusual ORFs which were likely to be bacterial in nature and which could not be combined with other obvious bacteriophage based contigs are not included in discussion of ORFs.

The extrapolation of the size of the genome by overlapping gene matches was carried out without the inclusion of those contigs suspected of being bacterial contamination. Without these contigs, the extrapolated genomic size was similar to that of the proposed size (Section 4.3.7). This is further evidence that these contigs may be the product of bacterial contamination. This conclusion depends on the accuracy of the calculation of size (4.3.7) as well as on the expectation that gene size in unsequenced regions will be the same as that of the gene to which it is matched. This is not guaranteed, or even likely in some gene regions. For any future primer directed approaches, where DNA yield is not as crucial, the use of DNase 1 in preparation of DNA would be prudent. This would eliminate the possibility of any PCR products being the result of non-specific binding to contaminating bacterial DNA.

The use of the theorised position of contigs based on matches to the *B. vietnamiensis* G4 genome, resulted in targeted sequencing which combined two contigs (A and B) to form a single large contig. Further primer design focussing on these hypothetical contig orientations may be of use in the combination of more contigs, should the time and funding be available. The matches to the *B. vietnamiensis* bacterial contigs can be theorised to be matches to a prophage present in the *B. vietnamiensis* genome. In fact, several of the putative *B. vietnamiensis* genes are reported as hypothetical bacteriophage genes. Of interest is the relationship between BupsΦ1, from a *B. pseudomallei* isolate originally isolated in North Queensland, and a possible prophage in *B. vietnamiensis*, a member of the *B. cepacia* complex isolated in Vietnam (Gillis *et al.*, 1995). Several of the ORFs identified on the BupsΦ1 genome also had strong matches to various bacteriophage from the *B. cepacia* complex. It is possible that BupsΦ1 was, at some time in the past, a bacteriophage infecting *B. cepacia* and changes to those genes responsible for recognition of cell wall structures resulted in a new host range in the related bacterium *B. pseudomallei*. Alternately there has been extensive horizontal exchange of DNA. In fact, the putative tail fibre gene, ORF D5, has a 48% identity match to the *B. cenocepacia* bacteriophage BcepMu. Casjens (2003) stated that; "...two phages of the same type are more likely to have a higher proportion of more closely related genes if they infect closely related hosts." and "...horizontal exchanges are common among the dsDNA tailed phages..." Should the genomic sequencing of BupsΦ1 be completed, it would be interesting to examine the relationship of the genome to *B. cepacia* and *B. pseudomallei* bacteriophage genomes already completed.

The genome of *B. pseudomallei* K96243 has been completely sequenced and three prophages identified within it (Holden *et al.*, 2004). BupsΦ1 had one very poor match (W2) with the non-induced prophage GI15 in K96243. The lack of matches compared to those of *B. vietnamiensis* seems to indicate that BupsΦ1 is dissimilar to at least the prophages of *B. pseudomallei* K96243.

6.4.2 Structural genes

The putative tail fibre gene appeared to have a mosaic structure, with matches to other tail fibre genes being partial and limited to particular regions or domains. This has been noted previously in large double stranded DNA bacteriophages (Sandmeier, 1994). Domain I matches were unrelated to bacteriophage of *Burkholderia* species, although *Ralstonia* species are included in the *Burkholderiaceae*. It is expected this region of the gene codes for the region of the tail fibres closest to the tail of the bacteriophage. The short unique region (domain II) may be a hinge. Domain III corresponds to a gene possibly involved with adhesion (COG3210) in *B. pseudomallei* as well as tail fibres of bacteriophage extracted from *B. cenocepacia*. This region is likely to be involved with binding to the *B. pseudomallei* cell wall. The similarities to these genes would be due to the similarities of cell wall structures of *Burkholderia* species. The matches are not identical, 48% (gp52) and 34% (COG3210) and this variability would be responsible for the variation in specificity of binding to isolates of *B. pseudomallei.*

Of the other structural proteins found on Myoviridae genomes, only the TerL, portal protein and head morphogenesis genes were putatively identified. Of these, TerL and the portal protein are considered to be the most conserved of the bacteriophage genes (Casjens, 2003). A partial ORF (Q1) for tail assembly was identified, although its specific purpose was not. The protease, scaffold, head/tail-joining proteins, tail shaft protein, tapemeasure protein, tail tip/baseplate proteins were not identified, though ORF Q1 may be one of the tail proteins. These proteins have been recorded as being very diverse, though conserved in order (Casjens, 2003) and may require experimentation to identify. Of note, ORF DU3 has a match to COG3299, an uncharacterised homolog of phage Mu protein gp47 in *E. coli*. Another homolog of phage Mu protein gp47, DVU1102, has been putatively identified as a baseplate assembly protein, so it is possible that, as this ORF is just upstream of the tail fibre protein, this is what ORF DU3 encodes for.

6.4.3 Replication

When the genome size was determined by pulse field gel electrophoresis (Section 4.3.7), no genomic ladder which would indicate end-annealed multimers (Summer *et al.*, 2004) was observed. This suggests that BupsΦ1 does not use *cos*-type DNA packaging.

ORFs putatively involved with DNA replication (Table 6.3) were examined to determine whether any further functions could be hypothesised. TerL, while considered to be structural, has a role in cleaving virion-length molecules from concatomeric replicating DNA (Casjens, 2003). Contig C carried only hypothetical and unidentified genes with the exception of C7. ORF C7 was identified as RNA polymerase, which would likely be used by the bacteriophage for transcribing late genes which have promoters not identified by the host's RNA polymerase. Of note are two inverted repeat regions, one of which has proposed short inverted repeat regions at either end and is likely to be the complete region. The other lies at one end of the contig and has an identical short region at the known end. This is indicative of transposition activity, as transposons typically have short, complementary repeat regions at either end (Hallet and Sherratt, 1997). In this case the whole region, where it is known, is complementary. It is possible that one of these regions has copied itself using a transposition mechanism and reciprocal recombination has then occurred to one copy, resulting in an inverted region bracketed by the inverted repeat regions. The value of this to the bacteriophage is unknown as all ORFs in these regions are hypothetical or have no match. However, it is interesting that ORF C7, a putative RNA polymerase sigma subunit, lies between these two regions, with some of C7 included in the second region. The large inverted repeat regions may have some conformational characteristics which could affect expression of the RNA polymerase and hence control replication of BupsΦ1.

Several possible transposases were identified by BLASTx homology (N2, N3, AB17, DU1). Of these, ORF AB17 (possible transposase) and ORF N2 (transposase) were

complete ORFs, while DU1 (transposase) and N3 (integrase or transposase) were partial ORFs.

It is tempting to speculate that BupsΦ1 uses replicative transposition to amplify the genome, integrating copies of the genome randomly throughout the host genome prior to excision and packaging. These types of bacteriophage have not been commonly described, but over 60 have been reported in *Pseudomonas*. The resultant bacteriophage genomes may well carry random chunks of host DNA at the ends of the genome. BcepMu of *B. cenocepacia* is also a transposable bacteriophage and typically carries 2000bp of host DNA at the right end of the genome (Summer *et al.*, 2004). Evidence to support this theory includes the multiple contigs of DNA which would not be expected to be carried on a bacteriophage genome (suspect contigs) as well as the variable virulence described in Chapter 5. A note of caution should be sounded regarding the use of the DNA evidence to support this theory. The genome has not been completed and variable regions of host DNA covalently bound to the end of the bacteriophage genome would be required to prove this case. In addition, the presence of a Holliday junction resolvase (ORF AB13) is indicative of λ like insertion (Hallet and Sherratt, 1997). Holliday junction resolvases are used to resolve the Holliday junction formed during integration of the genome.

Further examination identifies several ORFs putatively involved in lysogeny. ORF AB3 encodes a putative repressor gene related to SOS-response transcriptional repressors (RecA-mediated autopeptidases). Production of RecA is a repair mechanism generated by the host to respond to DNA damage and is used by lysogenic bacteriophage use to trigger entry into the lytic cycle. Such a repressor would maintain the lysogenic state until damage to the host's DNA induced lysis. In λ and other bacteriophage, Cro is produced after the repressor protein is removed. This stops the production of any more repressor protein, ensuring that the lysis continues. ORF AB4 is weakly matched to the Cro gene of the *S. typhimurium* bacteriophage ST104. Some of the genes then transcribed in λ include int and xis (integrase and excisase) which are used to excise the bacteriophage DNA and form

circular dsDNA. If BupsΦ1 replicates in a fashion similar to λ rather than Mu, the presence of these genes would be expected.

Partial ORF N3 matches an insertion element described as an integrase, although this is only 38 aa of a 286 aa gene. No excisase, as present in λ, was identified. In λ, integrase and excisase genes are positioned several genes to the left of the repressor cI. On contig AB, the first seven ORFs (AB1a, b, ca, c, d, e, AB2) have either no match or only a hypothetical match. The excisase gene may be present here, with Contig N placed previously. The integrase gene is generally adjacent to or very near the attachment site and this may be associated with ORF N1.

It is more likely that BupsΦ1 replicates in a similar fashion to λ than to Mu, as many of the genes necessary for λ type replication are present, while the evidence for Mu type replication is based on the presence of putative transposases

Cloning of lysin and holin genes can be problematic as they are often lethal to *E. coli* if expressed (Paul *et al.*, 2002). Bacteriophage of Gram negative organisms often have the following five genes in a cluster: holin, antiholin, endolysin and Rz/Rz1 equivalents (Wang *et al.*, 2000). Of these, only the endolysin is usually identifiable by homology searches (Summer *et al.*, 2004). The lysis genes are typically found in the same orientation, adjacent to and at either end of the structural cluster (Casjens, 2003), although this is not universal. As BupsΦ1 has been shown to lyse *B. pseudomallei*, an endolysin gene must be present on the genome, however no endolysin gene was identified by homology searches on the BupsΦ1 contigs. Use of ClustalX alignments and visualisation of chemo-physical relatedness did identify two ORFs (DU4 and AB15) as candidates for an endolysin gene. Of these, AB15 is more likely due to relative position on a contig containing structural ORFs. An attempt to quantify this relationship using phylogenetic tree analysis also showed AB15 to have a closer relationship than DU4 to the endolysin genes it was compared to. Examination of all ORFs to identify a holin gene, as well as direct examination of AB14, failed to identify any strong match to a holin sequence and indicated that AB14 shows no similarity to selected holin sequences. It has been noted that these

genes are often highly variable (Summer *et al.*, 2004), so this does not exclude this ORF from being a holin gene, but it cannot be used to support AB15 as an endolysin.

Sequential alignments with ClustalX were observed to result in variation of alignment and bootstrap values were well below the 70% score indicative of a real relationship. Analysis of many ORFs at once, where those ORFs represented unrelated genes, was problematic in that relationships may have been hidden in the larger picture. This was observed with bootstrap values, where trees calculated from smaller numbers of sequences gave different bootstrap values and different branching than trees consisting of all ORFs. The combination of ClustalX, GeneDoc and NJ plot were useful tools but required subjective interpretation. At best, this analysis has identified one or two ORFs that would be good candidates for targeted experimentation.

6.4.4 Virulence determinants

Several ORFs putatively encoded possible virulence determinants; ORF D2 encodes a possible oxidoreductase and varieties of these have been reported as virulence determinants (Ezraty *et al.*, 2005). The position of a virulence determinant near or within the tail region of the genome and next to a transposase, such as the case here, has also been described as common in bacteriophage (Boyd and Brussow, 2002). ORF E2 has been putatively identified as an anaerobic dehydrogenase. Anaerobic dehydrogenases are enzymes necessary for bacterial metabolism in anaerobic environments and this would give *B. pseudomallei* a survival advantage in infection of humans and animals, so may be a virulence determinant.

One of the virulence components of interest is that of an ADP-ribosylating exotoxin, although no ORF was identified by homology searching to match any recorded exotoxin. This is not unexpected as ADP-RTs typically have very different amino acid sequences but do have a conserved pattern of amino acids at the active site. Four ORFs were identified as having amino acid sequences with conserved regions similar to those ADP-RTs reported previously. Multiple ADP-ribosylating enzymes

have been reported on a single bacteriophage previously (Tiemann *et al.*, 2004). Bacteriophage T4 carries three ADP-ribosyltransferases. In the case of T4, these enzymes are not recorded as being toxins, but rather they modify host proteins such as RNA polymerase to control bacteriophage replication (Depping *et al.*, 2005). These enzymes also share the distinctive conserved amino acid sequence present on reported toxin ADP-RTs (Tiemann *et al.*, 2004).

Two of the four ORFs identified had matches to conserved amino acid patterns of group one ADP-RTs. These toxins attach to histidine on elongation factor 2 in sensitive cells, stopping protein synthesis. They include exotoxin A of *P. aeruginosa* and diphtheria toxin (Wozniak *et al.*, 1988). The third ORF had matches to conserved regions of group two ADP-RTs which ADP-ribosylate heterotrimeric G-proteins. These include cholera and pertussis toxin (Barth *et al.*, 1998). The fourth ORF had matches to group four exotoxins, which ADP-ribosylates G-actin and includes botulinum C2 toxin (Aktories *et al.*, 1986). Literature varies on the number and position of conserved amino acids in these toxins. Comparison of five exotoxins (ETA of *P. aeruginosa*, DT of *Corynebacterium diphtheriae*, S1 of *Bordetella pertussis*, CT of *Vibrio cholerae* and LTH of *E. coli*) has shown conservation of a required glutamic acid residue 124-125 residues downstream of a required histidine residue (Wozniak *et al.*, 1988) and followed by a tryptophan residue another four residues further on. Other literature (Takada *et al.*, 1995) confirms the requirement of the glutamic acid residue, but notes that the histidine upstream may be between 105 and 130aa away, and may in fact be an arginine. Takeda (1995) and others (Barth *et al.*, 1998; Han and Tainer, 2002; Oakey *et al.*, 2005) also record a second glutamic acid, two residues upstream from the first one, which is commonly conserved. This residue is not present in exotoxin A of *P. aeruginosa* (Gray *et al.*, 1984) or diphtheria toxin (Han and Tainer, 2002). No note is made of conservation of tryptophan but a region of high hydrophobicity and or aromaticity is described between the histidine/arginine and glutamic acid (Takada *et al.*, 1995). Group two to four toxins are also described as having an **STS** motif between the arginine and glutamic acid residues (Barth *et al.*, 1998).

In the case of those ORFs where no **STS** motif was present, sequences were examined for the presence of an area of hydrophobicity and those which did not have such a region were eliminated, leaving only ORFs AB10 and K1. All other expected residues and regions were apparent and the distances between residues was in-line with those recorded previously. The flexibility of the distances between some residues, as well as the number of ORFs that matched portions of these, mean the possibility of one or both of these ORFs fitting this pattern purely by chance increases. This random matching has been reported previously (Lax *et al.*, 1990; Lobet *et al.*, 1991).

In addition to the two ORFs showing similarities to group one ADP-RTs, two more possible ADP-RTs were identified matching to group two (AB12) and four (F1) ADP-RTs. Neither of these display the conserved glutamic acid two residues upstream of the glutamic acid residue required for activity, however they do display all the other conserved characteristics of the relevant ADP-RTs. Of these, ORF AB12 is a better candidate as ORF F1 is at the short limit of those ADP-RTs previously described. ORF AB12 matches very closely in length to the active region of cholera toxin and the putative VHML ADP-RT (Oakey *et al.*, 2005). The arginine residue of ORF AB12 is 122aa from the 5' end of the sequence and the end of the ORF is 198 aa after the glutamic acid residue. The active site is central to the ORF, with other regions both upstream and downstream. This is unlike that of the VHML ADP-RT, which lies at the C terminal end of the ORF and that of cholera toxin, which lies at the N terminal end of the ORF.

Several of these ORFs have homology matches to recorded genes. ORF AB10 has a 56% identity match to a PAPS reductase/sulfotransferase, AB12 has a weak partial match (38%) to a putative DNA binding protein, corresponding to part of the area of the active residues and K1 (a partial ORF) has a strong match (98% identity) with gp30 of *B. pseudomallei*, a hypothetical gene. PAPS reductase, to which AB10 was matched, is used in a step in the pathway of reduction of sulphur for the incorporation into cysteine and methionine (Williams *et al.*, 2002b). It catalyses the reduction of PAPS to sulphite in the presence of reduced thioredoxin (Schierova *et*

al., 2000). No PAPS reductase has been identified as an ADP-RT at this time. The closest relationship may be found in *Bordetella pertussis*. *B. pertussis* toxin has ADP-ribosylating activity and has, like other ADP-RTs, an absolutely required glutamic acid$_{(129)}$ (Locht *et al.*, 1989). Production of this toxin has been coupled to the presence of cysteine, with limited cysteine resulting in greater toxin production (Bogdan *et al.*, 2001). It has been shown that this is due to the drop in sulphate anions produced from cysteine, where lower cysteine levels means less sulphate is produced (Melton and Weiss, 1989). As PAPS reductase is involved with cysteine synthesis from sulphate (the reverse reaction) it would be involved in dropping the level of sulphate anions. If it were also to be an ADP-RT it could positively regulate its own production. The weak match of AB12 to a DNA binding molecule may indicate that the ADP-RT pattern identified in this region of the ORF could be involved in replication rather than production of a toxin. No speculation can be made about ORF K1 and F1 as these ORFs have no putative function associated with them.

Given the total of four possible ADP-RTs identified across three templates, it is highly likely that several or even all of these ORFs are not actually ADP-RTs. As matches to the templates have been previously shown to be by random chance in some cases, there is still doubt as to the identification of these ORFs as ADP-RTs. If the identification of ADP-ribosylating activity is correct in all four cases, some or all of them may in fact be involved in bacteriophage replication rather than toxin activity. Any identification of a virulence component should be treated with caution until further research is carried out to clarify the action of the proteins produced by these putative genes.

6.4.5 Conclusion

The analysis of the genome has only just begun. Complete sequencing of the genome will undoubtedly provide more information into the interaction of BupsΦ1 and its host. Only a few of the tools available to analyse bacteriophage DNA have been used thus far. For example, further examination of potential ADP-RTs could be carried out by generating secondary structures as all ADP-RTs examined in this

fashion have similar characteristics (Han and Tainer, 2002). Other methods of analysis include identification of origins of replication using cumulative GC-skew plots (Grigoriev, 1998; Tang *et al.*, 2002; Mehta *et al.*, 2004) and identification of tRNA like sequences (Tang *et al.*, 2002), identification of putative promoter and terminators (Mehta *et al.*, 2004) and examination of domains (http://au.expasy.org/prosite/). The only limit to "*in silico*" analysis seems to be time and computing power.

Confirmation of function can also be achieved experimentally. Northern blots could be used to determine the time of transcription of the various ORFs (Tang *et al.*, 2002). Recombinant production of gene products and mutagenesis analysis could also be used to experimentally confirm the putative functions of genes. While recombinant protein production is a time consuming step, functional assays on these ORF products can be simple. The determination of ADP ribosyl transferase activity *in vitro* is a simple assay, but hazardous to the user, involving extraction of elongation factor two (EF2) from wheatgerm and addition of the candidate protein and C14-NAD to this in 96 well plate format. After incubation, the reaction is stopped with trichloroacetic acid and the solution is collected on filter paper and washed to remove free C14-NAD. The radiosignal is determined and is indicative of incorporation of ADP-ribose in EF2 in the presence of C14-NAD; ADP ribosylating activity (Mohamed *et al.*, 1989b). As I had to develop a whole new procedure to undertake this assay and it was not central to the thesis, it was not taken any further at this time.

Most genomes recorded are not experimentally proven, or only partially so. The thorough experimental analysis of ORFs of any sequenced bacteriophage genome will aid in the expansion of knowledge of gene function and provide confirmed gene structures. A more thorough understanding of the role of the bacteriophage in its host is a worthy goal, given the relevant and varied role these organisms play in our ecosystem and in our health.

CHAPTER 7
INVESTIGATION OF BACTERIOPHAGE LYSIN

7.1 Introduction

Bacteriophage lytic enzymes quickly destroy the cell wall of the host bacterium to release progeny bacteriophage. Since such lytic enzymes specifically kill the species in which they were produced, they may represent an effective way to control pathogenic bacteria (Nelson *et al.*, 2001). These bacteriophage lysins hydrolyse the bacterial wall by cleaving bonds in the peptidoglycan layer and include muramidases, transglycosylases, L-alanine amidases and endopeptidases (Yoon *et al.*, 2001). Most bacteriophage lysins consist of two functional domains, the N-terminal domain which has the enzymatic activity as described above and the C-terminal domain which has cell wall binding capacity, allowing the N-terminal region to target its substrate (Loessner, 2005). In natural infection, for lysin to access the peptidoglycan layer, another protein called holin is required to first form lesions in the cytoplasmic membrane (Yoon *et al.*, 2001).

Lysins are of interest as they can also cleave the peptidoglycan layer if exposed to it "from without". That is, they can destroy cells with an external application without requiring holin to function. In addition, since infection of the bacteria is not required to destroy the bacteria, lysin can act where whole bacteriophage might be constrained by superinfection immunity mechanisms present in lysogens (Nesper *et al.*, 1999).

The specificity of lysins for particular bacteria or even strains of bacteria mean that they can target a single species or strain of bacteria (Schuch *et al.*, 2002). Extracted lysin, partially purified and characterised from group C streptococci (Fischetti *et al.*, 1971; Raina, 1981), has been shown to eliminate group A streptococci in the upper respiratory tract of mice, without killing of other microflora (Nelson *et al.*, 2001). Studies of an expression library of bacteriophage proteins from *Bacillus anthracis* identified a product, PlyG, capable of lysing *B. anthracis* and closely related organisms (Schuch *et al.*, 2002). Capsules were found not to block access of the protein to the cell.

From a molecular viewpoint, two experimental designs seem most popular in the identification of bacteriophage lysin (Table 7.1). The first involves production of an expression library, followed by screening of the products for lysin activity. The second involves sequencing of the genome, identification of genes which may be involved in lysin activity, amplification and expression of those genes and screening for lysin activity.

Table 7.1 Examples of experimental techniques used to identify bacteriophage lysin genes

bacterial host	bacteriophage	identification method	source
Lactobacillus plantarum	SC921	expression library, expression in *E. coli*	(Yoon *et al.*, 2001)
Lactobacillus bulgaris	mv1	expression library, expression in *E. coli*	(Boizet *et al.*, 1990)
Lactobacillus strain G1e	Φg1e	expression library, expression in *E. coli*	(Oki *et al.*, 1996)
Lactococcus lactis	ΦvML3	expression library, expression in *E. coli*	(Shearman *et al.*, 1989)
Lactobacillus gasseri	Φadh	genome sequencing, putative gene amplification, expression	(Henrich *et al.*, 1995)
Streptococcus thermophilus	ΦO1205	genome sequencing, putative gene amplification, expression	(Sheehan *et al.*, 1999)
Erwinia amylovora	ΦEa1h	genome sequencing, putative gene amplification, expression	(Kim *et al.*, 2004)

With the exception of *E. amylovora*, the examples in table 7.1 are all Gram-positive bacteria. These do not have an outer membrane layer covering the peptidoglycan layer as do Gram negative organisms and thus the lysin has easier access to the peptidoglycan layer. However, the outer membrane of Gram negative organisms is a semi-permeable layer, including channels (porins) allowing access to the peptidoglycan layer. Other substances which act on the peptidoglycan layer, such as β-lactams (Gould and MacKenzie, 1997; Nau and Eiffert, 2002), are effective against some Gram negative organisms. Given the success of lysis from without

with the Gram negative *E. amylovora*, it is possible that bacteriophage lysin may also be effective at lysing the cell from without in the case of *B. pseudomallei*.

There are three broad methods of production of bacteriophage lysin. The first is concentration and purification in natural hosts, the second is development of an expression library for screening, sequencing and experimentation and the last, sequencing of the genome to identify regions of interest by BLASTx searching, followed by expression of PCR products.

Extraction and purification of lysin from natural hosts has typically involved concentration steps, both prior to cracking of the bacteria (lysis from within) and/or after cracking. Chemicals may be added to stabilise and protect the lysin with procedures undertaken at 4°C wherever possible to slow degradation. Early studies (Fischetti *et al.*, 1971) to purify lysin from group C streptococcal bacteriophage involved a 652 fold concentration and purification of bacteria, prior to cracking and resuspension in media containing dithiothreitol (DTT) and DNase 1. EDTA was added after cracking and centrifugation carried out to remove cell debris. At this point tetrathionate was added to stabilise the enzyme and ammonium sulphate, chromatography and gel purification steps carried out. Lysin was stored at -51°C. This method was also used by Nelson *et al.* (2001) to extract and purify lysin for studies on treatment of upper respiratory tract infection.

In later work on the same lysin (Raina, 1981), the broth was not concentrated prior to cracking, but immediately afterwards, by chilling to 2°C with mercaptoethanol and EDTA. Two steps of ammonium sulphate precipitation, followed by gel purification and a final step of polyethylene glycol (PEG) concentration were conducted prior to storage of the enzyme at -75°C.

More recently, Hertwig *et al* (1997) purified and characterised lysin from bacteriophage P001 from *L. lactis*. Crude lysin was produced by concentration of broth to 1/15th its original volume in Tris buffer with glucose prior to bacterial cracking. Once cells lysed, DNase 1 and ribonuclease A (RNase A) were added and

the lysates centrifuged to remove bacteriophage and cell debris. The lysin was stored at -74°C prior to examination for its ability to lyse killed cells. Further purification was carried out with addition of trifluoroacetic acid (TFA), anion exchange, cation exchange and HPLC chromatography.

Production of lysin via molecular methods have been typically carried out in one of two ways. In the first, gene expression libraries from the bacteriophage were constructed and screened for expression of lysin. Examples of this include ΦvML3 of *L. lactis* (Shearman *et al.*, 1989) and SC921 of *L. plantarum* (Yoon *et al.*, 2001). In both these cases, bacteriophage DNA was digested with restriction endonucleases and inserted into expression vectors. Individual clones were grown, concentrated and treated with chloroform to break open the cells. Cell debris was removed and the cell-free supernatants inserted into wells on an agar plate pre-seeded with bacteria sensitive to the bacteriophage examined. Plates were examined after overnight incubation for clearing around the wells indicative of lysin. Further molecular analysis of the selected clones and of the cell-free extracts was then undertaken to characterise the lysin and lysin gene. An adaptation of this method was used to identify the lysin gene of ϒ phage of *B. anthracis* (Schuch *et al.*, 2002). In this case, the clone colonies were grown overnight on glass plates, had their cell walls permeabilised by chloroform vapour and then had soft agar containing indicator bacteria poured over the colonies. Plates were incubated and clones containing lysin genes were identified by the presence of clear lytic zones. The bacteriophage mv1 lysin gene was identified using clones of *Escherichia coli* directly spotted on a lawn of indicator bacteria (Boizet *et al.*, 1990). Several clones that produced zones of lysis also had very low viability. One difficulty with screening for lysin genes by the method of expression libraries is that, if the holin gene is included in the expression vector along with the lysin gene, lysis from within can occur and lower the viability of the *E. coli* host. This may have also been seen in one clone during the work of Shearman *et al.* (1989).

The second approach to lysin production requires some previous knowledge of the lysin gene and involves development of primers to produce the lysin gene as a PCR

product, followed by in frame insertion into an expression vector. The lysin of ΦO1205 of *S. thermophilus* was produced by designing primers for expected lysin genes based on published gene sequence of ΦO1205 (Sheehan *et al.*, 1999) as was the lysin of ΦEa1h of *E. amylovora* (Kim *et al.*, 2004). In both cases PCR products were inserted into an expression vector and grown in *E. coli.* This was then sonicated to release lysin and the lysin tested against indicator bacteria. No treatment of the *E. amylovora* was carried out to aid access to the peptidoglycan layer prior to effective lysis from without.

Treatment and removal of *B. pseudomallei* from the environment is difficult, hence new therapeutic substances are highly desirable. Using the literature, techniques for production of lysin from the natural host will be adapted. The resultant products will be tested to see whether they have any effect on cells of *B. pseudomallei* which can be attributed to a proteinaceous lysin rather than bacteriophage. Any product with a lytic effect will be further purified and characterised.

7.2 Materials and Methods

7.2.1 Amplification of *B. pseudomallei* in broth

A colony of *B. pseudomallei* isolate #4 was inoculated into 100 ml LB broth and incubated with shaking overnight (100rpm, 37°C). Ten ml of this broth was inoculated into 850ml-950ml LB and incubated with shaking (100rpm, 37°C) until an $O.D._{600nm}$ of 0.1 was obtained. Fifty ml was removed for use as a negative control and a Miles-Misra assay carried out. Fifty ml was removed for use as a positive control and BupsΦ1 was added to a ratio of 10:1 to 1:1 cfu:pfu to act as the positive control. The broths for both the positive and negative control were then incubated with shaking and the remaining 750ml-850ml of broth was retained at 37°C . Between 30 minutes and one hour later, the 750ml-850ml of broth was diluted with prewarmed LB media to adjust $O.D._{600nm}$ back to 0.1 and BupsΦ1 was added at the same ratio of 10:1 to 1:1 cfu:pfu as the positive control broth. All broths were further incubated with shaking (100rpm, 37°C) until lysis of bacterial cells was seen in the positive control broth. Lysis was identified as a clearing of the broth with "strings" of bacteria visible.

7.2.2 Concentration of broth and addition of chemicals

DTT was added to broths to a concentration of 0.5mM after which they were centrifuged (10 000*g*, 5 minutes, 4°C) either ten minutes before lysis was expected or when lysis was first evident, whichever came first. The pellet was resuspended to 1/15th of the original volume in tris-glucose buffer (Appendix 1). Where lysis did not occur at the expected time, chloroform was added at 0.1% to 1% of solution volume. EDTA was added to a concentration of 5mM either when lysis was expected to occur or when lysis was observed. After lysis, the solution was centrifuged (9000*g*, ten minutes). DNase 1 and RNase were added to the supernatant at a concentration of 170µg/ml and 1.7mg/ml respectively and the supernatant incubated at 37°C for 15 minutes. This was centrifuged (9000*g*, ten minutes, 4°C), the supernatant filtered through a 0.45 µm filter (Millex® -GS, Millipore, Ireland) and stored at -74°C until used.

7.2.3 Production of indicator bacteria

7.2.3.1 Live bacteria

One ml of an overnight LB broth of *B. pseudomallei* strain #4 was added to one ml of LB soft top agar. Two ml of the resultant solution was poured onto a LB agar plate and allowed to set.

7.2.3.2 Killed bacteria

A 100ml overnight broth of bacteria (isolate #4) was concentrated 100x by centrifugation (10 000*g*, 5 minutes) and heated at 95°C for 20 minutes in a water bath. The water bath was turned off and the bacteria retained in the water bath for a further 20 minutes. To test if the bacteria were dead, a loop of treated bacteria was placed in LB broth and incubating overnight with shaking (100rpm, 37°C). Bacterial cells were considered dead if no bacterial growth was evident. For plate use, 0.5 ml of the concentrated killed cells was added to one ml of soft top agar and the solution poured onto a LB agar plate. For assays using optical density, the $O.D._{600nm}$ of cells was adjusted to 0.6 and a 1:10 mix of lysin extract to dead cells was examined. Tris-glucose buffer was used under the same conditions, instead of lysis extract, as a negative control.

7.2.4 Analysis of extracts for lytic activity

7.2.4.1 Use of indicator bacteria

Extracts were tested for their ability to lyse cells using both killed and live cells. Six µl of extract (7.2.2) from #4 broth infected with BupsΦ1 (7.2.1) was placed on LB soft top plates containing bacterial cells and the plates incubated at 37°C. Plates were observed at ten minutes, then at 30 minute intervals for four hours, and finally after overnight incubation. Killed cells and extract were mixed in cuvettes at the ratio noted previously (7.2.3.2) and the optical density observed immediately, with observation continuing for ten minutes and then at the intervals described above.

7.2.4.2 Fractionation of extracts

Extracts were fractionated at sizes of >30kDa, 30-10kDa and <10kDa with Centricon© centrifugal filters (Millipore, USA). To a 30kDa Centricon, 1.5 ml of extract was added and the unit centrifuged (4500*g*, 30 minutes, 4°C). Retentate was collected by inversion of the Centricon and centrifuging into a collection unit (200*g*, two minutes, 4°C). Where no fluid retentate was visible, 50µl of PBS was added to the Centricon and the fluid pipetted several times prior to inversion and centrifugation. The filtrate was added to a 10kDa Centricon and the process repeated, with an initial centrifugation being carried out for one hour, as per the manufacturer's instructions. In this case both retentate and filtrate were collected. When not being centrifuged, all fractions were stored on ice. All fractions were collected and absorbance at 280nm was determined prior to storage at -74°C. Six µl of each was spotted on a soft top agar plate containing killed cells.

7.2.4.3 Production of bacterial supernatant by bead beating

To produce a negative control sample of cell contents, 50 ml of an overnight culture of *B. pseudomallei* #4 which had not been exposed to BupsΦ1 was centrifuged (8500*g*, ten minutes) and the pellet resuspended in three ml of Tris-glucose buffer with 300µl of 0.5M EDTA. This was added to a two ml free standing screw cap microtube with o-ring (520-GRDS-Q, QSP, USA) containing one ml of 0.1mm glass beads (Biospec Products Inc., USA) such that no air gap was present when the cap was screwed on. This was exposed to three pulses each of 45 seconds duration in a

Mini-Beadbeater™ (Biospec Products inc., USA) with one minute on ice between each pulse. The sample was then centrifuged (10 000*g*, ten minutes) and the supernatant collected and stored at -20°C until needed.

7.2.4.4 SDS polyacrylamide gel electrophoresis analysis

Thirty µl of each extract and fractionated extract was run on a 4-12% Nu-PAGE™ BT 1.0 SDS polyacrylamide gel electrophoresis (SDS-PAGE) (Invitrogen, Australia) after which it was stained with Gradipure (Gradipore, Australia) according to the manufacturer's instructions. Following staining, the gel was destained with 6% acetic acid and dried for long term storage using a Gel Drying Kit (Promega) according to the manufacturer's instructions.

7.3 Results

7.3.1 Lysis

In spite of similar bacteriophage to bacteria ratios, smaller volumes of media (50ml) showed a greater disposition to lyse than larger volumes (750-850ml) (Table 7.2). It was observed that lysis took longer in 750-850ml samples and was less complete. This was the case both where the sample was centrifuged and made into a small volume prior to lysis and when lysis was permitted to occur prior to centrifugation to remove cell debris. Addition of chloroform did not trigger lysis. After addition of EDTA, concentrated samples lysed and became viscous. Addition of DNase 1 and RNase, reduced viscosity, indicating that most viscosity was caused by nucleic acid.

7.3.2 Examination of extracts

Bacteriophage extracts that had been stored at 4°C in excess of two weeks were used to test whether killed cells could be lysed by active bacteriophage. These extracts caused clearing on soft top agar containing live cells, but not on soft top agar containing killed cells.

All lysed extracts were tested using soft top agar containing killed cells and absorbance at $O.D._{600nm}$ analysis using killed cells (7.2.4.1). In no case did the agar plates clear or the optical density drop. Plating of extracts on soft top agar

containing live bacteria resulted in clearance indicative of the presence of bacteriophage.

Lysis extracts were examined using SDS PAGE. Multiple bands were present and bands were similar to those produced from a bead-beated supernatant of *B. pseudomallei* #4 which had not been infected with bacteriophage (Figure 7.1). Several bands were detected in an extract which had not shown any evident lysis (Figure 7.1, lane 4). Two of these bands were the correct size to be DNase 1 (~31 kDa) and RNase A (~13.7 kDa), which had been added to this extract.

Table 7.2 Conditions of individual lysis experiments and their outcomes. Time to concentration dependant on lysis of control sample. Time was counted from point of adding bacteriophage to broth. Experiments 1 and 2 were pilot experiments, did not use controls and were concentrated at arbitrary times. Experiment 4 included a control (B) of the same volume as the sample (A), made at the same time. This was concentrated after lysis was evident. Complete lysis means the broth was observed to be completely cleared of bacteria. Visible lysis means the presence of "strings" of lysed bacteria were present. Minimal lysis means some clumps of bacteria were present, but not enough to form strings.

Experiment	volume (ml)	time to concentration (hour:minute)	chemicals added after concentration in Tris-glucose	time to lysis	conditions of lysis	observed lytic activity of extract
1	850	0:25	none	-	none	none
2	850	1:00	none	-	none	none
3	800	1:30	DNase1, RNase A	1:45	visible, viscous	none
4(A)	400	1:00	none	3:00	visible	none
(B)	400	3:50	none	3:00	lysis visible prior to concentration, no enhancement of lysis after this	none
5	50	2:10	none	3:10	complete	none
	850	2:10	chloroform	3:10	visible, chloroform did not appear to accelerate or improve lysis	none
6	50	1:30	DNase1, RNase A	2:30	complete, very viscous	none
	800	1:30	none	4:30	minimal	none
7	850	1:00	none	-	none	none
	850	1:30	none	-	none	none
	850	2:00	none	5:00	minimal	none

Table 7.2 continued.

Experiment	volume (ml)	time to concentration (hour:minute)	chemicals added after concentration in Tris-glucose	time to lysis	conditions of lysis	observed lytic activity of extract
	850	2:30	none	-	none	none
	850	5:00	none	5:00	complete	none
8	800	0:40	DTT, chloroform, EDTA	3:45	chloroform caused no change. Lysis visible 15 minutes after addition of EDTA	none
	800	1:15	DTT, chloroform, EDTA	3:45	chloroform caused no change. Lysis visible 15 minutes after addition of EDTA	none
9(A)	50	1:50	EDTA, DNase1, RNase A	2:00	visible, viscous	none
(B)	50	1:50	EDTA, DTT	2:00	visible	none
	800	1:50	EDTA, DTT	3:30	limited	none

Fractions concentrated using Centricons™ were compared to the original sample using $O.D._{280nm}$ readings (Biophotometer; Eppendorf, Germany) (Table 7.3). Sample volumes were corrected for concentration where possible prior to SDS PAGE. Based on $O.D._{280nm}$ readings, retention in a 30kDa filter resulted in approximately five times the concentration. Examination of the gel indicated all bands were present in this fraction (Figure 7.2). The 30-10kDa fraction was also concentrated by a factor of approximately ten. The low volume retained meant $O.D._{280nm}$ readings and SDS PAGE were carried out using a diluted fraction, created by addition of PBS to the membrane and recovery as described above (Section 7.2.4.2). Proteins of less than 33 kDa, based on the PAGERuler protein ladder, were recovered in this fraction. No protein fractions were of high enough concentration to be visible in the <10kDa fraction. About 5% volume was lost in the total concentration protocol. No fraction caused lysis of killed cells in spite of an increase in concentration.

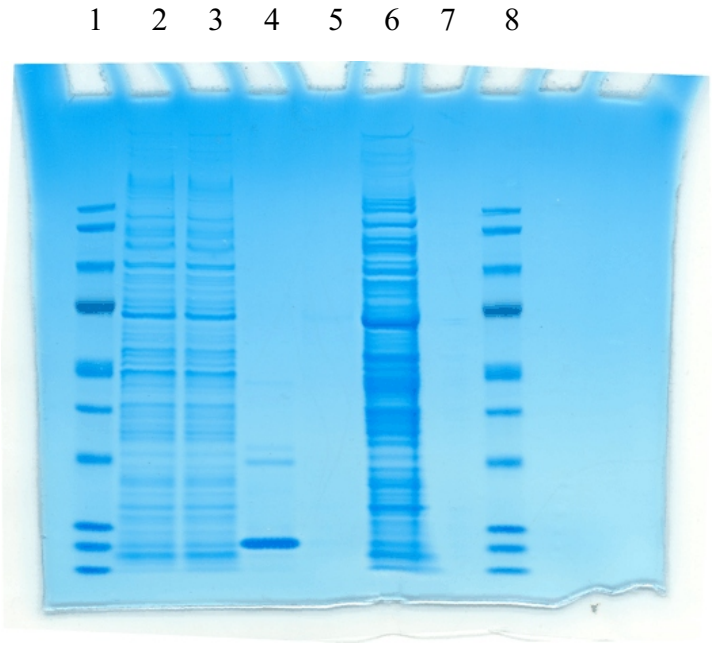


Figure 7.1 SDS-PAGE of products of experimental lysis (taken from experiment 9A-Table 7.2), lane 1 and 8; PAGERuler™ Prestained Protein Ladder (Fermentas, Germany) {mwt (kDa) 170, 130, 100, 72, 55, 40, 33, 24, 17, 11}, lane 2; lysin extract without ten mM DTT, lane 3; lysin extract with ten mM DTT, lane 4; non-lysed broth supernatant including DTT, DNase 1 and RNase A, lane 5; Tris-glucose buffer only, Lane 6; supernatant of *B. pseudomallei* #4 (not infected with bacteriophage) after bead beating, Lane 7; no sample.

Table 7.3 Volumes recovered and optical density (280nm) readings of lysis extract and fractions concentrated through 30kDa and 10kDa Centricons™. Samples include retentate of 30kDa membrane (>30kDa), filtrate of 30kDa membrane which was retained by the 10kDa membrane (<30kDa>10kDa) and filtrate of the 10kDa membrane

Sample	Volume used or recovered (µl)	Optical density (280nm) (corrected for dilutions)
lysis extract	1500	1.644
>30kDa fraction	190	7.115
<30kDa >10kDa fraction	5	17.75
<10kDa fraction	1220	0.541

Figure 7.2 SDS-PAGE of concentrated extracts of lysis from experiment 9A, , lane 1 and 6 ; PAGERuler™ Prestained Protein Ladder {mwt (kDa) 170, 130, 100, 72, 55, 40, 33, 24, 17, 11}, lane 2; fraction >30kDa, lane 3; fraction <30kDa>10kDa, lane 4; fraction <10kDa, lane 5; extract after lysis, prior to concentration.

7.4 Discussion

Fischetti *et al.* (1971) and Nelson *et al.* (2001) used DTT when extracting lysin, both while concentrating and cracking the bacteria and again later to reactivate lysin. DTT breaks disulphide bonds and maintains SH groups in a reduced state by

preventing oxidation (Sigma catalogue). Raina (1981) used 2-mercaptoethanol, which has a similar activity to DTT, whilst Hertwig *et al.* (1997) used neither. However, no lysis was found to occur with crude lysin extracted from BupsΦ1 irrespective of whether DTT was added or not.

EDTA and chloroform have been used to permeabilise the cell wall, permitting recombinant lysin to escape and aiding in cracking of bacterial cells (Henrich *et al.*, 1995; Yoon *et al.*, 2001) and to make Gram negative organisms susceptible to lysis from without (Morita *et al.*, 2001; Loessner, 2005). In this case, chloroform was not found to aid cracking. However, EDTA was successfully used to cause mass lysis of concentrated bacterial broths. When EDTA was not added, lysis was observed to be less complete (Table 7.2). Irrespective of whether or not EDTA was added, no lysis on killed cells was observed, confirming that lysis from without did not occur with this Gram negative bacteria.

Reported lysin sizes range between 20kDa (vML3 of *L. lactis*) (Shearman *et al.*, 1994), to 98kDa (P22 of *Salmonella typhimurium*) (Henrich *et al.*, 1995). SDS PAGE showed the presence of multiple bands within this range, however none clearly differed from the bands produced when uninfected *B. pseudomallei* #4 bacteria was treated with a bead beater to lyse the cells compared to those that lysed due to bacteriophage. The efficient lysis of the bacterial broths indicates lysin should be present, but identification of a single band is not possible. A more effective method of observing lysin on SDS-PAGE gels would possibly be to identify the lysin gene by sequencing and insertion into an expression vector. Lysin could then be overproduced and purified for study. Success of this process is dependant on not lysing the *E. coli* cell before high levels of lysin are produced. Examples in literature indicate this method is feasible (Shearman *et al.*, 1994; Henrich *et al.*, 1995; Sheehan *et al.*, 1999) and failure is usually due to the simultaneous production of holin.

In spite of successfully causing lysis in concentrated bacterial broths, with typical nucleic acid viscosity as described by Hertwig *et al.* (1997), no lysis of killed cells

was observed, either in soft top agars or by a drop in $O.D._{600nm}$. As crude extracts and concentrates did not kill *Burkholderia* cells, no further purification was carried out. It is thought that either the extraction technique failed to concentrate sufficient active lysin, or the lysin could not lyse *B. pseudomallei* cells externally. The first possibility is less likely as concentration by fractionation also failed to cause lysis. The variation between experiments in the timing of the concentration step, due to variation in individual control samples, also did not identify an optimum time for production of lysin.

The possibility that bacteriophage lysin does not work from without in the case of *B. pseudomallei* must be considered further. *B. pseudomallei* is known to possess a capsule (Puthucheary *et al.*, 1996) which may be protective, however there is some evidence that capsules do not impede lysis from without (Henrich *et al.*, 1995).

A recent review article (Loessner, 2005) states that lysis from without is not possible with Gram negative organisms due to the presence of the outer membrane of the cell wall. *B. pseudomallei* is a Gram negative organism so lysis from without may not be possible.

Permeability properties of the outer membrane control access of solutes to the periplasm and inner membrane. Transit of hydrophilic solutes across the outer membrane is accomplished by a battery of proteins forming transmembrane channels of varying degrees of specificity (Osborn and Wu, 1980). There are three functional categories of transmembrane channel. The first are porins, allowing passive, nonspecific diffusion of solutes across the membrane. Only hydrophilic solutes smaller than the exclusion limit specific for a given Gram negative bacterium can pass the outer membrane (Benz, 1988), with selectivity due to the size of the hydrated radius of solute versus the size of the pore. The second are pore forming proteins allowing specific permeation of oligosaccharides of maltose and nucleosides. The third allows uptake of large solutes of low concentration with specificity due to receptors interacting with high affinity, for example receptors for

B12 and iron siderophores (Osborn and Wu, 1980). These last two types of channel are designed for specific uptake and not relevant to the possible uptake of lysins.

Porins, the first category, permit β-lactams to access the peptidoglycan layer. β-lactams are hydrophilic and usually diffuse through porins (Gotoh *et al.*, 1994a). Work with β-lactams in *E. coli* (Nikaido *et al.*, 1983) showed that hydrophilic β-lactams diffused through porins with the maximum size of the β-lactams tested of a molecular weight of 527 (5kDa). The pore of the porin was determined to be 1.2nm across (Nikaido and Rosenberg, 1981). The largest recorded exclusion limit (molecular weight above which the molecule cannot pass through) for a porin is 6kDa, found in *P. aeruginosa* (Hancock *et al.*, 1979; Bellido *et al.*, 1992). This is much smaller than any recorded endolysin. While there are examples of specific transfer pores in *Burkholderia* species, there are none recorded of a size large enough to permit lysin access (Gotoh *et al.*, 1994a; Gotoh *et al.*, 1994b; Tsujimoto *et al.*, 1997; Adewoye *et al.*, 1998). The smallest lysin would still be too large to diffuse through the largest recorded porin. Even though lysins are generally hydrophilic (Loessner, 2005), the method used by β-lactams to penetrate to the peptidoglycan layer is not available to endolysins.

As expected, almost all examples in the literature of lysis from without are of Gram positive organisms. One exception is the recombinant lysin of ΦEa1h of *E. amylovora* (Kim *et al.*, 2004). In this case, cell lysate from sonicated *E. coli* cells was found to directly induce lysis on lawns of indicator cells and cell growth in broths was inhibited when cell lysates were added. The overproduction of lysin in a recombinant system may be the reason lysis was seen by Kim *et al*. (2004). Further exceptions exist as stated by Morita *et al*. (2001). While it is generally the case that peptidoglycan degrading enzymes are limited in activity against Gram negative bacteria, the endolysin of the bacteriophage of *Bacillus amyloliquefaciens* effectively decreased viability of the Gram negative organism *Pseudomonas aeruginosa.* In addition to this it causes lysis from within when recombinantly produced by *E. coli*. This lysin was shown to contain a hydrophobic region which

may have been penetrating the outer membrane and allowing access by the catalytic domain of the endolysin to the peptidoglycan layer of the Gram negative bacteria.

The lack of success in causing lysis from without using concentrated lysate in this study may indicate that the endolysin BupsΦ1 of *B. pseudomallei* does not contain hydrophobic characteristics like that of the endolysin of the bacteriophage of *B. amyloliquefaciens*.

Given the compelling evidence that without addition of chemicals to disrupt the outer membrane, lysin will not function "from without" in *B. pseudomallei*, it is unlikely that production of recombinant lysin would be of any practical therapeutic use. This chapter confirms, in the specific case of *B. pseudomallei*, the general statement that 'lysis from without does not occur in Gram negative organisms'.

CHAPTER 8
GENERAL DISCUSSION

Burkholderia pseudomallei was first identified in 1911 (Whitmore, 1913) and bacteriophage specific to it were first identified in 1956 (Leclerc and Sureau, 1956). *B. pseudomallei* is the causative agent of melioidosis, a disease of humans and animals with a high mortality rate. The title of this thesis asked the question; Are bacteriophage of *B. pseudomallei* friend or foe? That is, are they responsible in some way for the virulence of *B. pseudomallei*, or can they be used therapeutically?

In many cases bacteriophage have been found to be responsible for virulence of bacteria (Miller, 2001; Boyd, 2004). This role for bacteriophage with respect to *B. pseudomallei* was examined in this study. Bacteriophage have also been used as treatments for bacterial infections as both entire bacteriophage (Smith *et al.*, 1987; Huff *et al.*, 2004; Jikia *et al.*, 2005) and as protein products such as endolysin (Nelson *et al.*, 2001). In this study, endolysin was examined from lysogenic bacteriophage to determine whether it would be of use.

This study identified combinations of techniques for highly specific determination of bacteriophage infection in *B. pseudomallei* as well as expanding on the data on bacteriophage infecting *B. pseudomallei*. This study also examined the role of one of these bacteriophage (BupsΦ1) in virulence in both an animal model and by analysis of the genome of the bacteriophage. Furthermore, the use of bacteriophage endolysin as an external therapy for melioidosis was examined and rejected.

The traditional method of screening and amplifying *B. pseudomallei* for bacteriophage involves the use of *B. mallei* as a sensitive host (Denisov and Kapliev, 1991; Denisov and Kapliev, 1995; Woods *et al.*, 2002). *B. mallei* is highly infectious to humans, causing a life threatening disease (Smith, 1957) and has been known to cause infections via laboratory contamination (Howe and Miller, 1947; Srinivasan *et al.*, 2001). As such, this study examined alternate methods for identification of lysogenic bacteriophage. A simple technique for screening large numbers of isolates for identification of isolates of interest was developed and it was

determined that the use of a mitomycin C method in conjunction with either a plaque assay or restriction digest assay resulted in confirmation of the presence of bacteriophage with an accuracy of 96.77%. Using either a plaque or restriction digest assay had the added value of allowing some characterisation of the bacteriophage, either spotting pattern or restriction cutting pattern respectively. The optimisation of various techniques of amplification of bacteriophage and extraction of DNA with respect to bacteriophage of *B. pseudomallei* will be of use in further examination of bacteriophage of other *B. pseudomallei* strains.

Bacteriophage of *B. pseudomallei* have previously been identified as belonging to the families *Myoviridae* and *Syphoviridae* (Denisov and Kapliev, 1995). The isolate further characterised in this study was identified as a member of the *Myoviridae*, with a genome 55.1 kb long. Of interest, the bacteriophage isolated from strain C4 was of the family *Myoviridae* but had an unusually long tail. This may be due to a tape-measure gene more similar to those in the family *Syphoviridae*. This isolate would be of interest to examine further in a separate study.

Lysogenic bacteriophage insert their genome into the bacterial chromosome and once inserted, genes on the bacteriophage can be expressed, conferring new metabolic traits on the host. The process of altering properties of the host bacterial cell on establishment of lysogeny is known as phage conversion (Acheson *et al.*, 1998). Bacteriophage may encode virulence factors which can increase the virulence of their host. Several bacterial - bacteriophage systems have been analysed in which bacteriophage extracted from a virulent strain of bacteria have been inserted into a strain of lower virulence and caused an upregulation of virulence (Munro *et al.*, 2003). To date no examination of alteration of virulence due to bacteriophage infection in an animal model has been reported for *B. pseudomallei*.

This study is the first examination of the role of bacteriophage in *B. pseudomallei* isolates, using the BALB/c mouse model of Leakey *et al.* (1998). It was shown that bacteriophage could infect some of the bacterial isolates selected and alter colonial morphology in some cases. There were also indications of variable phenotypic

changes. There was no clear relationship between bacteriophage infection and virulence, with only one assay resulting in a significant increase in virulence. It is possible that bacteriophage stability was the cause of this variation in virulence. A bacteriophage that was lysogenically stable would permit expression of virulence determinants and would not cause significant lysis of the bacteria. If the relationship was less stable, as appeared to be the case when comparing NCTC13178 and NAFC, more of the bacteria would lyse due to bacteriophage entry into the lytic lifecycle. This would lower virulence through lowering levels of bacteria as well as limiting the possible expression of any bacteriophage encoded virulence. The variation in virulence seen in the NAFC-BupsΦ1 system may have been due to genome insertion position affecting overall bacteriophage stability and hence both effective bacterial dose and expression of virulence determinants.

The selection of BupsΦ1 and host system NAFC may have been a poor choice, given the possibility of other complicating factors. However, BupsΦ1 infection of other strains was also unsuccessful. If bacteriophage do play a role in virulence of *B. pseudomallei*, that role is not a straight forward one and other factors may be involved. This is the case in virulence of *Vibrio cholerae*, where the presence of a toxin co-regulated pilus, believed to be encoded by a potential second bacteriophage (Karaolis *et al.*, 1999) is necessary for infection with the bacteriophage encoding the cholera toxin gene (Levin and Tauxe, 1996). In addition, recent reviews of virulence of *V. cholerae* have shown the expression of toxin genes encoded on CTXΦ to be promoted by a gene encoded on the satellite bacteriophage RS1 (Davis and Waldor, 2003). Amplification and examination of other bacteriophage in more *B. pseudomallei* isolates may be of value, as other bacterial-bacteriophage systems may show clearer virulence effects. Determination of the insertion site may also shed some light onto the effect the bacteriophage are having in the host.

Examination of the genetic sequence of bacteriophage can aid in understanding bacteriophage-host interactions including virulence (Uetz *et al.*, 2004). As there was no clear relationship between bacteriophage infection and virulence in this study, bacteriophage BupsΦ1 was examined at a genetic level. Only two isolated

bacteriophage of *B. pseudomallei* have been completely sequenced (NCBI Entrez genome; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Pager&DB=genome), although sequencing of bacteriophage from the related organisms *B. thailandensis* (Woods *et al.*, 2002) and *B. cenocepacia* (Summer *et al.*, 2004; Seed and Dennis, 2005) has been carried out and several bacteriophage of *B. pseudomallei* have been sequenced as prophage (Holden *et al.*, 2004), so this sequencing adds significantly to the amount of data available.

In this study, 95% of the genome of BupsΦ1 (based on determination of genome size by pulse field gel electrophoresis) was sequenced and the open reading frames analysed using BLASTx. There were frequent matches to open reading frames of a putative prophage in the genome of *B. vietnamiensis* and these open reading frames were typically in the same order in both BupsΦ1 and the reported prophage. The number of matches to the prophage relative to other related bacteriophage may indicate a common origin, however this may be an artifact due to the paucity of sequence data of *Burkholderia* bacteriophage genomes. Sequencing of this and further *B. pseudomallei* bacteriophage genomes will aid in the understanding of the evolutionary development of these systems.

Examination of the sequence data identified putative genes involved in structural development, DNA replication, lysogeny and virulence. Alternate tools were also used to data mine the genomic information in a search for genes of interest. These identified possible ADP-ribosyltransferase genes and endolysin genes which would not have otherwise been identified.

Experimental analysis would be necessary to confirm the function of these open reading frames. Many of the reported genes that were used to identify putative function in the BLASTx analysis were putative themselves, with no experimental evidence to confirm function. Future work examining the open reading frames of this bacteriophage at an experimental level would be of value for further sequence analysis.

The presence of putative virulence genes is of interest as no clear role in virulence was identified in the animal model. Should these putative genes prove to be legitimate, it would further reinforce the idea that the role BupsΦ1 is playing in virulence is complex and dependant on other factors that have not yet been identified.

In addition to playing a role in virulence, bacteriophage can play a role in the treatment of bacterial infection. Bacteriophage cocktails have been used to successfully treat antibiotic resistant infections (Jikia *et al.*, 2005) but a lysogenic bacteriophage which may increase the virulence of its host could not be used in this manner. Alternately, the endolysin protein produced by bacteriophage to lyse the bacteria from within can be used from without (Nelson *et al.*, 2001). While it has been stated that endolysin cannot be used on Gram negative organisms (Loessner, 2005), exceptions exist. These include the recombinant lysin of ΦEa1h of *E. amylovora* (Kim *et al.*, 2004) and the endolysin of the bacteriophage of *B. amyloliquefaciens*, which effectively decreased viability of the Gram negative organism *P. aeruginosa* (Morita *et al.*, 2001).

Examination of the endolysin of BupsΦ1 in this study showed that endolysin will not function "from without" in *B. pseudomallei*. This study confirms, in the specific case of *B. pseudomallei*, the general statement that 'lysis from without does not occur in Gram negative organisms'. As such, it is unlikely that production of recombinant lysin would be of any practical therapeutic use. The next step in examination of bacteriophage for therapeutic work would involve the isolation and amplification of bacteriophage of *B. pseudomallei* which are obligatorily lytic.

This thesis asked the question; Are bacteriophage of *B. pseudomallei* friend or foe? We can now say that lysogenic bacteriophage of *B. pseudomallei* cannot be used as a 'friend'. However, this study did not identify or examine obligatorily lytic bacteriophage, so this area remains open to investigation. The role of bacteriophage as a 'foe' is more complex. There are indications of possible virulence determinants as well as a role for instability of the bacteriophage. There was no simple answer to

whether the presence or absence of bacteriophage related to virulence from the initial screening of *Burkholderia* isolates. Examination of a single bacteriophage (BupsΦ1) did not clearly show that bacteriophage were responsible for virulence. It is possible that some lysogenic bacteriophage of *B. pseudomallei* (including possibly this one) carry virulence determinants and sequencing is ongoing (Wiersinga *et al.*, 2006). Bacteriophage of *B. pseudomallei* may be our 'foe', but this work has raised more questions for study than it has been able to answer.

In conclusion, this work provides a foundation for the examination of bacteriophage of *B. pseudomallei* using methods of lower risk than those traditionally used. It increases the quanta of data available on what isolates are infected with bacteriophage and their cross-spotting patterns as well as on the gene sequences found in *B. pseudomallei* bacteriophage genomes. This work also provides the first examination of the role of bacteriophage in virulence of *B. pseudomallei* in a mouse model. This indicates any role of virulence may be more complex than that of a simple virulence gene encoded on the bacteriophage. This study also eliminates the possibility of the use of endolysin in the treatment of *B. pseudomallei* infection.

REFERENCES

Acheson, D.W., Reidl, J., Zhang, X., Keusch, G.T., Mekalanos, J.J. and Waldor, M.K. (1998) In vivo transduction with shiga toxin 1-encoding phage. *Infect Immun* **66**: 4496-8

Ackermann, H.W. (1998) Tailed bacteriophages: the Order Caudovirales. *Adv Virus Res* **51**: 135-201

Ackermann, H.W. (2003) Bacteriophage observations and evolution. *Res Microbiol* **154**: 245-51

Ackermann, H.W., Audurier, A., Berthiaume, L., Jones, L.A., Mayo, J.A. and Vidaver, A.K. (1978) Guidelines for bacteriophage characterization. *Adv Virus Res* **23**: 1-24

Ackermann, H.W., DuBow, M.S., Jarvis, A.W., Jones, L.A., Krylov, V.N., Maniloff, J., Rocourt, J., Safferman, R.S., Schneider, J., Seldin, L. and et al. (1992) The species concept and its application to tailed phages. *Arch Virol* **124**: 69-82

Adewoye, L.O., Tschetter, L., O'Neil, J. and Worobec, E.A. (1998) Channel specificity and secondary structure of the glucose-inducible porins of Pseudomonas spp. *J Bioenerg Biomembr* **30**: 257-67

Aktories, K., Barmann, M., Ohishi, I., Tsuyama, S., Jakobs, K.H. and Habermann, E. (1986) Botulinum C2 toxin ADP-ribosylates actin. *Nature* **322**: 390-2

Alberts, B., Bray, D., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (1998) Essential Cell Biology; An Introduction to the Molecular Biology of the Cell, 630pp pp, edited by B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter. Garland, New York.

Allison, G.E., Angeles, D., Tran-Dinh, N. and Verma, N.K. (2002) Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of *Shigella flexneri*. *J Bacteriol* **184**: 1974-87

Anne, J., Verheyen, P., Volckaert, G. and Eyssen, H. (1985) A restriction endonuclease map of *Streptomyces* phage VWB. *Mol Gen Genet* **200**: 506-7

Anne, J., Wohlleben, W., Burkardt, H.J., Springer, R. and Puhler, A. (1984) Morphological and molecular characterization of several actinophages isolated from soil which lyse *Streptomyces cattleya* or *S. venezuelae*. *J Gen Microbiol* **130 (Pt 10)**: 2639-49

Anon. (1921) The Phenomenon of d'Herelle. *The Lancet* **198**: 1173-1174

Austin, B., Pride, A.C. and Rhodie, G.A. (2003) Association of a bacteriophage with virulence in *Vibrio harveyi*. *J Fish Dis* **26**: 55-8

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, S.C., Seidman, J.G., Smith, J.A. and Struhl, K. (1992) Short Protocols in Molecular Biology. 2 ed, edited by F.M. Ausubel, R. Brent, R.E. Kingston, S.C. Moore, J.G. Seidman, J.A. Smith and K. Struhl. Green Publishing, New York.

Baess, I. (1971) Report on a pseudolysogenic *mycobacterium* and a review of the literature concerning pseudolysogeny. *Acta Pathol Microbiol Scand [B] Microbiol Immunol* **79**: 428-34

Barksdale, L. and Arden, S.B. (1974) Persisting bacteriophage infections, lysogeny, and phage conversions. *Annu Rev Microbiol* **28**: 265-99

Barnes, J.L. and Ketheesan, N. (2005) Route of infection in melioidosis. *Emerg Infect Dis* **11**: 638-9

Barreiro, V. and Haggard-Ljungquist, E. (1992) Attachment sites for bacteriophage P2 on the *Escherichia coli* chromosome: DNA sequences, localization on the physical map, and detection of a P2-like remnant in *E. coli* K-12 derivatives. *J Bacteriol* **174**: 4086-93

Barth, H., Preiss, J.C., Hofmann, F. and Aktories, K. (1998) Characterization of the catalytic site of the ADP-ribosyltransferase *Clostridium botulinum* C2 toxin by site-directed mutagenesis. *J Biol Chem* **273**: 29506-11

Baselski, V.S., Medina, R.A. and Parker, C.D. (1979) In vivo and in vitro characterization of virulence-deficient mutants of *Vibrio cholerae*. *Infect Immun* **24**: 111-6

Bellido, F., Martin, N.L., Siehnel, R.J. and Hancock, R.E. (1992) Reevaluation, using intact cells, of the exclusion limit and role of porin OprF in *Pseudomonas aeruginosa* outer membrane permeability. *J Bacteriol* **174**: 5196-203

Bensing, B.A., Siboo, I.R. and Sullam, P.M. (2001) Proteins PblA and PblB of *Streptococcus mitis*, which promote binding to human platelets, are encoded within a lysogenic bacteriophage. *Infect Immun* **69**: 6186-92

Benson, S.D., Bamford, J.K., Bamford, D.H. and Burnett, R.M. (2004) Does common architecture reveal a viral lineage spanning all three domains of life? *Mol Cell* **16**: 673-85

Bentley, R. (2005) The development of penicillin: genesis of a famous antibiotic. *Perspect Biol Med* **48**: 444-52

Benz, R. (1988) Structure and function of porins from Gram negative bacteria. *Annu Rev Microbiol* **42**: 359-93

Benzer, S. (1955) Fine structure of a genetic region in bacteriophage. *Proc Natl Acad Sci U S A* **41**: 344-354

Bernhardt, T.G., Wang, I.N., Struck, D.K. and Young, R. (2002) Breaking free: "protein antibiotics" and phage lysis. *Res Microbiol* **153**: 493-501

Bernstein, C. (1981) Deoxyribonucleic acid repair in bacteriophage. *Microbiol Rev* **45**: 72-98

Bertani, G. and Nice, S.J. (1954) Studies on lysogenesis. II. The effect of temperature on the lysogenization of *Shigella dysenteriae* with phage P1. *J Bacteriol* **67**: 202-9

Bertani, G. and Weigle, J.J. (1953) Host controlled variation in bacterial viruses. *J Bacteriol* **65**: 113-21

Bertani, L.E. (1971) Stabilization of P2 tandem double lysogens by int mutations in the prophage. *Virology* **46**: 426-36

Betley, M.J., Miller, V.L. and Mekalanos, J.J. (1986) Genetics of bacterial enterotoxins. *Annu Rev Microbiol* **40**: 577-605

Beutin, L., Stroeher, U.H. and Manning, P.A. (1993) Isolation of enterohemolysin (Ehly2)-associated sequences encoded on temperate phages of *Escherichia coli*. *Gene* **132**: 95-9

Billing, E. (1969) Isolation, Growth and Preservation of Bacteriophages. In: J.R. Norris and D.W. Ribbons (Eds), Methods in Microbiology, pp. 315-330. Vol. 3B. Academic Press, London.

Blakely, G.W. (2004) Smarter than the average phage. *Mol Microbiol* **54**: 851-4
Bogdan, J.A., Nazario-Larrieu, J., Sarwar, J., Alexander, P. and Blake, M.S. (2001) *Bordetella pertussis* autoregulates pertussis toxin production through the metabolism of cysteine. *Infect Immun* **69**: 6823-30
Boizet, B., Lahbib-Mansais, Y., Dupont, L., Ritzenthaler, P. and Mata, M. (1990) Cloning, expression and sequence analysis of an endolysin-encoding gene of *Lactobacillus bulgaricus* bacteriophage mv1. *Gene* **94**: 61-7
Bordenstein, S.R. and Wernegreen, J.J. (2004) Bacteriophage flux in endosymbionts (wolbachia): infection frequency, lateral transfer, and recombination rates. *Mol Biol Evol* **21**: 1981-91
Bossi, P., Tegnell, A., Baka, A., van Loock, F., Hendriks, J., Werner, A., Maidhof, H. and Gouvras, G. (2004) Bichat guidelines for the clinical management of glanders and melioidosis and bioterrorism-related glanders and melioidosis. *Euro Surveill* **9**
Bourniquel, A.A. and Bickle, T.A. (2002) Complex restriction enzymes: NTP-driven molecular motors. *Biochimie* **84**: 1047-59
Boyd, E.F. (2004) Bacteriophages and Bacterial Virulence. In: E. Kutter and A. Sulakvelidze (Eds), Bacteriophages: Biology and Applications, pp. 223-266. CRC Press.
Boyd, E.F. and Brussow, H. (2002) Common themes among bacteriophage-encoded virulence factors and diversity among the bacteriophages involved. *Trends Microbiol* **10**: 521-9
Boyd, J.S.K. (1953) Mutation in bacterial viruses. In: P. Fildes and W.E. Van Heyningen (Eds), The nature of virus multiplication: second symposium of the Society for General Microbiology held at Oxford University, April 1952, pp. 320. Cambridge University Press, Cambridge.
Bradley, D.E. (1967) Ultrastructure of bacteriophage and bacteriocins. *Bacteriol Rev* **31**: 230-314
Braid, M.D., Silhavy, J.L., Kitts, C.L., Cano, R.J. and Howe, M.M. (2004) Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*. *J Bacteriol* **186**: 6560-74
Broudy, T.B., Pancholi, V. and Fischetti, V.A. (2001) Induction of lysogenic bacteriophage and phage-associated toxin from group A streptococci during coculture with human pharyngeal cells. *Infect Immun* **69**: 1440-3
Bryant, J.L., Jr. and King, J. (1984) DNA injection proteins are targets of acridine-sensitized photoinactivation of bacteriophage P22. *J Mol Biol* **180**: 837-63
Calendar, R. (1970) The regulation of phage development. *Annu Rev Microbiol* **24**: 241-96
Carlson, P. and Seppanen, M. (2000) Melioidosis presenting as urinary tract infection in a previously healthy tourist. *Scand J Infect Dis* **32**: 92-3
Carrera, M.R., Kaufmann, G.F., Mee, J.M., Meijler, M.M., Koob, G.F. and Janda, K.D. (2004) Treating cocaine addiction with viruses. *Proc Natl Acad Sci U S A* **101**: 10416-21
Casjens, S. (2003) Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol* **49**: 277-300

Chanishvili, N., Chanishvili, T., Tediashvili, M. and Barrow, P.A. (2001) Phages and their application against drug-resistant bacteria. *J. Chem. Tech. Biotech* **76**: 689-699

Chaowagul, W., White, N.J., Dance, D.A., Wattanagoon, Y., Naigowit, P., Davis, T.M., Looareesuwan, S. and Pitakwatchara, N. (1989) Melioidosis: a major cause of community-acquired septicemia in northeastern Thailand. *J Infect Dis* **159**: 890-9

Charbit, A. and Hofnung, M. (1985) Isolation of different bacteriophages using the LamB protein for adsorption on *Escherichia coli* K-12. *J Virol* **53**: 667-71

Cheetham, B.F. and Katz, M.E. (1995) A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. *Mol Microbiol* **18**: 201-8

Chen, Y., Golding, I., Sawai, S., Guo, L. and Cox, E.C. (2005) Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. *PLoS Biol* **3**: e229

Cheng, A.C., Dance, D.A. and Currie, B.J. (2005a) Bioterrorism, Glanders and melioidosis. *Euro Surveill* **10**

Cheng, Q., Nelson, D., Zhu, S. and Fischetti, V.A. (2005b) Removal of group B *streptococci* colonizing the vagina and oropharynx of mice with a bacteriophage lytic enzyme. *Antimicrob Agents Chemother* **49**: 111-7

Chopin, A., Bolotin, A., Sorokin, A., Ehrlich, S.D. and Chopin, M. (2001) Analysis of six prophages in *Lactococcus lactis* IL1403: different genetic structure of temperate and virulent phage populations. *Nucleic Acids Res* **29**: 644-51

Chua, K.L., Chan, Y.Y. and Gan, Y.H. (2003) Flagella are virulence determinants of *Burkholderia pseudomallei*. *Infect Immun* **71**: 1622-9

Clapper, B., Tu, A.H., Elgavish, A. and Dybvig, K. (2004) The vir gene of bacteriophage MAV1 confers resistance to phage infection on *Mycoplasma arthritidis*. *J Bacteriol* **186**: 5715-20

Clark, C.A., Beltrame, J. and Manning, P.A. (1991) The oac gene encoding a lipopolysaccharide O-antigen acetylase maps adjacent to the integrase-encoding gene on the genome of *Shigella flexneri* bacteriophage Sf6. *Gene* **107**: 43-52

Cluzel, P.J., Serio, J. and Accolas, J.P. (1987) Interactions of *Lactobacillus bulgaricus* Temperate Bacteriophage 0448 with Host Strains. *Appl Environ Microbiol* **53**: 1850-1854

Coleman, D.C., Sullivan, D.J., Russell, R.J., Arbuthnott, J.P., Carey, B.F. and Pomeroy, H.M. (1989) *Staphylococcus aureus* bacteriophages mediating the simultaneous lysogenic conversion of beta-lysin, staphylokinase and enterotoxin A: molecular mechanism of triple conversion. *J Gen Microbiol* **135 (Pt 6)**: 1679-97

Collier, R.J. and Kandel, J. (1971) Structure and activity of diphtheria toxin. I. Thiol-dependent dissociation of a fraction of toxin into enzymically active and inactive fragments. *J Biol Chem* **246**: 1496-503

Cottew, G.S., Sutherland, A.K. and Meehan, J.F. (1952) Melioidosis in sheep in Queensland. Description of an outbreak. *Aust Vet J* **28**: 113-121

Crawford, J.T. and Goldberg, E.B. (1980) The function of tail fibers in triggering baseplate expansion of bacteriophage T4. *J Mol Biol* **139**: 679-90

Crowther, R.A., Lenk, E.V., Kikuchi, Y. and King, J. (1977) Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4. *J Mol Biol* **116**: 489-523

Currie, B., Smith-Vaughan, H., Golledge, C., Buller, N., Sriprakash, K.S. and Kemp, D.J. (1994) *Pseudomonas pseudomallei* isolates collected over 25 years from a non-tropical endemic focus show clonality on the basis of ribotyping. *Epidemiol Infect* **113**: 307-12

Currie, B.J., Fisher, D.A., Howard, D.M., Burrow, J.N., Selvanayagam, S., Snelling, P.L., Anstey, N.M. and Mayo, M.J. (2000) The epidemiology of melioidosis in Australia and Papua New Guinea. *Acta Trop* **74**: 121-7

Currie, B.J., Jacups, S.P., Cheng, A.C., Fisher, D.A., Anstey, N.M., Huffam, S.E. and Krause, V.L. (2004) Melioidosis epidemiology and risk factors from a prospective whole-population study in northern Australia. *Trop Med Int Health* **9**: 1167-74

Dabrowska, K., Switala-Jelen, K., Opolski, A., Weber-Dabrowska, B. and Gorski, A. (2005) Bacteriophage penetration in vertebrates. *J Appl Microbiol* **98**: 7-13

Dai, J.H., Lee, Y.S. and Wong, H.C. (1992) Effects of iron limitation on production of a siderophore, outer membrane proteins, and hemolysin and on hydrophobicity, cell adherence, and lethality for mice of *Vibrio parahaemolyticus*. *Infect Immun* **60**: 2952-6

Dance, D.A. (1991) Melioidosis: the tip of the iceberg? *Clin Microbiol Rev* **4**: 52-60

Dance, D.A. (2000) Melioidosis as an emerging global problem. *Acta Trop* **74**: 115-9

Dance, D.A., Smith, M.D., Aucken, H. and Pitt, T.I. (1999) Imported melioidosis in England and Wales. *The Lancet* **353**: 208

Davis, B.M. and Waldor, M.K. (2003) Filamentous phages linked to virulence of *Vibrio cholerae*. *Curr Opin Microbiol* **6**: 35-42

de Ruyter, P.G., Kuipers, O.P., Meijer, W.C. and de Vos, W.M. (1997) Food-grade controlled lysis of *Lactococcus lactis* for accelerated cheese ripening. *Nat Biotechnol* **15**: 976-9

Denisov, II and Kapliev, V.I. (1991) The level of spontaneous phage production and sensitivity to melioidosis phages of museum cultures of *Pseudomonas pseudomallei*. *Mikrobiol Zh* **53**: 66-70

Denisov, II and Kapliev, V.I. (1995) [The isolation and characteristics of cloned strains of *Pseudomonas pseudomallei* phages]. *Mikrobiol Z* **57**: 53-6

Depping, R., Lohaus, C., Meyer, H.E. and Ruger, W. (2005) The mono-ADP-ribosyltransferases Alt and ModB of bacteriophage T4: target proteins identified. *Biochem Biophys Res Commun* **335**: 1217-23

DeShazer, D. (2004) Genomic diversity of *Burkholderia pseudomallei* clinical isolates: subtractive hybridization reveals a *Burkholderia mallei*-specific prophage in B. pseudomallei 1026b. *J Bacteriol* **186**: 3938-50

Desmarchelier, P.M., Dance, D.A., Chaowagul, W., Suputtamongkol, Y., White, N.J. and Pitt, T.L. (1993) Relationships among *Pseudomonas pseudomallei* isolates from patients with recurrent melioidosis. *J Clin Microbiol* **31**: 1592-6

Diaz, E., Lopez, R. and Garcia, J.L. (1992) EJ-1, a temperate bacteriophage of *Streptococcus pneumoniae* with a Myoviridae morphotype. *J Bacteriol* **174**: 5516-25

Douglas, J. (1975) Bacteriophages. Chapman and Hall Ltd.

Dubaquie, Y. and Lowman, H.B. (1999) Total alanine-scanning mutagenesis of insulin-like growth factor I (IGF-I) identifies differential binding epitopes for IGFBP-1 and IGFBP-3. *Biochemistry* **38**: 6386-96

Duckworth, D.H. (1976) "Who discovered bacteriophage?" *Bacteriol Rev* **40**: 793-802

Eklund, M.W., Poysky, F.T., Peterson, M.E. and Meyers, J.A. (1976) Relationship of bacteriophages to alpha toxin production in *Clostridium novyi* types A and B. *Infect Immun* **14**: 793-803

Eklund, M.W., Poysky, F.T., Reed, S.M. and Smith, C.A. (1971) Bacteriophage and the toxigenicity of *Clostridium botulinum* type C. *Science* **172**: 480-2

Ezraty, B., Aussel, L. and Barras, F. (2005) Methionine sulfoxide reductases in prokaryotes. *Biochim Biophys Acta* **1703**: 221-9

Faruque, S.M., Asadulghani, Saha, M.N., Alim, A.R., Albert, M.J., Islam, K.M. and Mekalanos, J.J. (1998) Analysis of clinical and environmental strains of nontoxigenic *Vibrio cholerae* for susceptibility to CTXPhi: molecular basis for origination of new strains with epidemic potential. *Infect Immun* **66**: 5819-25

Fischetti, V.A., Gotschlich, E.C. and Bernheimer, A.W. (1971) Purification and physical properties of group C streptococcal phage-associated lysin. *J Exp Med* **133**: 1105-17

Fogelman, I., Davey, V., Ochs, H.D., Elashoff, M., Feinberg, M.B., Mican, J., Siegel, J.P., Sneller, M. and Lane, H.C. (2000) Evaluation of CD4+ T cell function In vivo in HIV-infected patients as measured by bacteriophage phiX174 immunization. *J Infect Dis* **182**: 435-41

Freeman, V.J. (1951) Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J Bacteriol* **61**: 675-88

Freeman, V.J. and Morse, I.U. (1952) Further observations on the change to virulence of bacteriophage-infected a virulent strains of *Corynebacterium diphtheria. J Bacteriol* **63**: 407-14

Frenkel, D. and Solomon, B. (2002) Filamentous phage as vector-mediated antibody delivery to the brain. *Proc Natl Acad Sci U S A* **99**: 5675-9

Fuhrman, J.A. (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541-8

Fuhrman, J.A. and Schwalbach, M. (2003) Viral influence on aquatic bacterial communities. *Biol Bull* **204**: 192-5

Gamage, S.D., Patton, A.K., Hanson, J.F. and Weiss, A.A. (2004) Diversity and host range of Shiga toxin-encoding phage. *Infect Immun* **72**: 7131-9

Gamage, S.D., Strasser, J.E., Chalk, C.L. and Weiss, A.A. (2003) Nonpathogenic *Escherichia coli* can contribute to the production of Shiga toxin. *Infect Immun* **71**: 3107-15

Gathogo, E.W., Waugh, A.C., Peric-Concha, N., Redpath, M.B. and Long, P.F. (2004) Rapid detection of 'rare' *actinomycetes* in environmental samples. *Biotechnol Lett* **26**: 897-900

Gehring, K., Charbit, A., Brissaud, E. and Hofnung, M. (1987) Bacteriophage lambda receptor site on the *Escherichia coli* K-12 LamB protein. *J Bacteriol* **169**: 2103-6

Gertman, E., White, B.N., Berry, D. and Kropinski, A.M. (1986) IS222, a new insertion element associated with the genome of *Pseudomonas aeruginosa*. *J Bacteriol* **166**: 1134-6

Gibbs, K.A., Isaac, D.D., Xu, J., Hendrix, R.W., Silhavy, T.J. and Theriot, J.A. (2004) Complex spatial distribution and dynamics of an abundant *Escherichia coli* outer membrane protein, LamB. *Mol Microbiol* **53**: 1771-83

Gillis, M., Van, T.V., Bardin, R., Goor, M., Hebbar, P., Willems, A., Segers, P., Kersters, K., Heulin, t. and Fernandes, M.P. (1995) Polyphasic taxonomy in the Genus *Burkholderia* leading to an emended description of the genus and Proposition of *Burkholderia vietnamiensis* sp. nov. for N2-fixing isolates from rice in Vietnam. *Int. J. Syst. Bacteriol.* **45**: 274-289

Gorski, A. and Weber-Dabrowska, B. (2005) The potential role of endogenous bacteriophages in controlling invading pathogens. *Cell Mol Life Sci* **62**: 511-9

Gotoh, N., Nagino, K., Wada, K., Tsujimoto, H. and Nishino, T. (1994a) *Burkholderia* (formerly *Pseudomonas*) *cepacia* porin is an oligomer composed of two component proteins. *Microbiology* **140 (Pt 12)**: 3285-91

Gotoh, N., White, N.J., Chaowagul, W. and Woods, D.E. (1994b) Isolation and characterization of the outer-membrane proteins of *Burkholderia* (*Pseudomonas*) *pseudomallei*. *Microbiology* **140 (Pt 4)**: 797-805

Gould, I.M. and MacKenzie, F.M. (1997) The response of Enterobacteriaceae to beta-lactam antibiotics--'round forms, filaments and the root of all evil'. *J Antimicrob Chemother* **40**: 495-9

Govan, B. and Ketheesan, N. (2004) Exposure to *Burkholderia pseudomallei* induces cell-mediated immunity in healthy individuals. *Clin Microbiol Infect* **10**: 585-7

Gray, G.L., Smith, D.H., Baldridge, J.S., Harkins, R.N., Vasil, M.L., Chen, E.Y. and Heyneker, H.L. (1984) Cloning, nucleotide sequence, and expression in *Escherichia coli* of the exotoxin A structural gene of *Pseudomonas aeruginosa*. *Proc Natl Acad Sci U S A* **81**: 2645-9

Greenfield, L., Bjorn, M.J., Horn, G., Fong, D., Buck, G.A., Collier, R.J. and Kaplan, D.A. (1983) Nucleotide sequence of the structural gene for diphtheria toxin carried by corynebacteriophage beta. *Proc Natl Acad Sci USA* **80**: 6853-7

Grigoriev, A. (1998) Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res* **26**: 2286-90

Groman, N. and Laird, W. (1977) Bacteriophage production by doubly lysogenic *Corynebacterium diphtheriae*. *J Virol* **23**: 592-8

Hagens, S., Habel, A., von Ahsen, U., von Gabain, A. and Blasi, U. (2004) Therapy of experimental *pseudomonas* infections with a nonreplicating genetically modified phage. *Antimicrob Agents Chemother* **48**: 3817-22

Hallet, B. and Sherratt, D.J. (1997) Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol Rev* **21**: 157-78

Han, S. and Tainer, J.A. (2002) The ARTT motif and a unified structural understanding of substrate recognition in ADP-ribosylating bacterial toxins and eukaryotic ADP-ribosyltransferases. *Int J Med Microbiol* **291**: 523-9

Hancock, R.E., Decad, G.M. and Nikaido, H. (1979) Identification of the protein producing transmembrane diffusion pores in the outer membrane of *Pseudomonas aeruginosa* PA01. *Biochim Biophys Acta* **554**: 323-31

Hardt, W.D., Urlaub, H. and Galan, J.E. (1998) A substrate of the centisome 63 type III protein secretion system of *Salmonella typhimurium* is encoded by a cryptic bacteriophage. *Proc Natl Acad Sci U S A* **95**: 2574-9

Hayashi, T., Baba, T., Matsumoto, H. and Terawaki, Y. (1990) Phage-conversion of cytotoxin production in *Pseudomonas aeruginosa*. *Mol Microbiol* **4**: 1703-9

Hayashi, T., Matsumoto, H., Ohnishi, M. and Terawaki, Y. (1993) Molecular analysis of a cytotoxin-converting phage, phi CTX, of *Pseudomonas aeruginosa*: structure of the attP-cos-ctx region and integration into the serine tRNA gene. *Mol Microbiol* **7**: 657-67

Hayes, W. (1964) The Genetics of Bacteria and their Viruses: Studies in basic genetics and molecular biology. Oxford: Blackwell scientific.

Hazan, R. and Engelberg-Kulka, H. (2004) *Escherichia coli* mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Mol Genet Genomics* **272**: 227-34

Henrich, B., Binishofer, B. and Blasi, U. (1995) Primary structure and functional analysis of the lysis genes of *Lactobacillus gasseri* bacteriophage phi adh. *J Bacteriol* **177**: 723-32

Hershey, A.D. and Chase, M. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* **36**: 39-56

Hertveldt, K., Lavigne, R., Pleteneva, E., Sernova, N., Kurochkina, L., Korchevskii, R., Robben, J., Mesyanzhinov, V., Krylov, V.N. and Volckaert, G. (2005) Genome comparison of *Pseudomonas aeruginosa* large phages. *J Mol Biol* **354**: 536-45

Hertwig, S., Bockelmann, W. and Teuber, M. (1997) Purification and characterization of the lytic activity induced by the prolate-headed bacteriophage P001 in *Lactococcus lactis*. *J Appl Microbiol* **82**: 233-9

Hitch, G., Pratten, J. and Taylor, P.W. (2004) Isolation of bacteriophages from the oral cavity. *Lett Appl Microbiol* **39**: 215-9

Holden, M.T., Titball, R.W., Peacock, S.J., Cerdeno-Tarraga, A.M., Atkins, T., Crossman, L.C., Pitt, T., Churcher, C., Mungall, K., Bentley, S.D., Sebaihia, M., Thomson, N.R., Bason, N., Beacham, I.R., Brooks, K., Brown, K.A., Brown, N.F., Challis, G.L., Cherevach, I., Chillingworth, T., Cronin, A., Crossett, B., Davis, P., DeShazer, D., Feltwell, T., Fraser, A., Hance, Z., Hauser, H., Holroyd, S., Jagels, K., Keith, K.E., Maddison, M., Moule, S., Price, C., Quail, M.A., Rabbinowitsch, E., Rutherford, K., Sanders, M., Simmonds, M., Songsivilai, S., Stevens, K., Tumapa, S., Vesaratchavest, M., Whitehead, S., Yeats, C., Barrell, B.G., Oyston, P.C. and Parkhill, J. (2004) Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc Natl Acad Sci U S A* **101**: 14240-5

Hoogenboom, H.R., Griffiths, A.D., Johnson, K.S., Chiswell, D.J., Hudson, P. and Winter, G. (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic Acids Res* **19**: 4133-7

Hooper, I., Woods, W.H. and Egan, B. (1981) Coliphage 186 Replication is delayed when the host cell is UV irradiated before infection. *J Virol* **40**: 341-9

Howe, C. and Miller, W.R. (1947) Human Glanders: Report of Six Cases. *Ann Intern Med* **26**: 93-115

Hsu, M.T. and Davidson, N. (1972) Structure of inserted bacteriophage Mu-1 DNA and physical mapping of bacterial genes by Mu-1 DNA insertion. *Proc Natl Acad Sci U S A* **69**: 2823-7

Huff, W.E., Huff, G.R., Rath, N.C., Balog, J.M. and Donoghue, A.M. (2004) Therapeutic efficacy of bacteriophage and baytril (enrofloxacin) individually and in combination to treat colibacillosis in broilers. *Poult Sci* **83**: 1944-7

Ilag, L.L., Tuech, J.K., Beisner, L.A., Sumrada, R.A. and Incardona, N.L. (1993) Role of DNA-protein interactions in bacteriophage phi X174 DNA injection. *J Mol Biol* **229**: 671-84

Jacob, F. and Fuerst, C.R. (1958) The mechanism of lysis by phage studied with defective lysogenic bacteria. *J Gen Microbiol* **18**: 518-26

Jacob, F. and Wollman, E.L. (1959) Lysogeny. In: F.M. Burnett and W.M. Stanley (Eds), The Viruses. Vol. 2. 3 vols. Academic Press, New York.

Jiang, S.C. and Paul, J.H. (1998) Significance of Lysogeny in the Marine Environment: Studies with Isolates and a Model of Lysogenic Phage Production. *Microb Ecol* **35**: 235-43

Jikia, D., Chkhaidze, N., Imedashvili, E., Mgaloblishvili, I., Tsitlanadze, G., Katsarava, R., Glenn Morris, J., Jr. and Sulakvelidze, A. (2005) The use of a novel biodegradable preparation capable of the sustained release of bacteriophages and ciprofloxacin, in the complex treatment of multidrug-resistant Staphylococcus aureus-infected local radiation injuries caused by exposure to Sr90. *Clin Exp Dermatol* **30**: 23-6

Johnson, L.P., Tomai, M.A. and Schlievert, P.M. (1986) Bacteriophage involvement in group A streptococcal pyrogenic exotoxin A production. *J Bacteriol* **166**: 623-7

Jones, A.L., Beveridge, T.J. and Woods, D.E. (1996) Intracellular survival of *Burkholderia pseudomallei*. *Infect Immun* **64**: 782-90

Kalantri, S., Pai, M., Pascopella, L., Riley, L. and Reingold, A. (2005) Bacteriophage- based tests for the detection of *Mycobacterium tuberculosis* in clinical specimens: a systematic review and meta- analysis. *BMC Infect Dis* **5**: 59

Karam, J.D. (2005) Bacteriophages: The viruses for all seasons of molecular biology. *Virol J* **2**: 19

Karaolis, D.K., Somara, S., Maneval, D.R., Jr., Johnson, J.A. and Kaper, J.B. (1999) A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in *cholera* bacteria. *Nature* **399**: 375-9

Karcher, A.M., Zaman, A., Brewis, C. and Fahmy, T. (2000) Neck lumps: expect the unexpected. *Lancet* **355**: 1070

Katsura, I. and Hendrix, R.W. (1984) Length determination in bacteriophage lambda tails. *Cell* **39**: 691-8

Keen, N., Staskawicz, B., Mekalanos, J., Ausubel, F. and Cook, R.J. (2000) Pathogens and hosts: the dance is the same, the couples are different. *Proc Natl Acad Sci U S A* **97**: 8752-3

Ketheesan, N., Barnes, J.L., Ulett, G.C., VanGessel, H.J., Norton, R.E., Hirst, R.G. and LaBrooy, J.T. (2002) Demonstration of a cell-mediated immune response in melioidosis. *J Infect Dis* **186**: 286-9

Kidambi, S.P., Booth, M.G., Kokjohn, T.A. and Miller, R.V. (1996) recA-dependence of the response of *Pseudomonas aeruginosa* to UVA and UVB irradiation. *Microbiology* **142 (Pt 4)**: 1033-40

Kilic, A.O., Pavlova, S.I., Alpay, S., Kilic, S.S. and Tao, L. (2001) Comparative study of vaginal *Lactobacillus* phages isolated from women in the United States and Turkey: prevalence, morphology, host range, and DNA homology. *Clin Diagn Lab Immunol* **8**: 31-9

Kilic, A.O., Pavlova, S.I., Ma, W.G. and Tao, L. (1996) Analysis of *Lactobacillus* phages and bacteriocins in American dairy products and characterization of a phage isolated from yogurt. *Appl Environ Microbiol* **62**: 2111-6

Kim, W.S., Salm, H. and Geider, K. (2004) Expression of bacteriophage phiEa1h lysozyme in *Escherichia coli* and its activity in growth inhibition of *Erwinia amylovora. Microbiology* **150**: 2707-14

Kingston, C.W. (1971) Chronic or latent melioidosis. *Med J Aust* **2**: 618-21

Kokjohn, T.A. (1989) Transduction: mechanism and potential for gene transfer in the environment. In: S.B. Levy and R.V. Miller (Eds), Gene Transfer in the Environment, pp. 73-97. McGraw-Hill Publishing Company, New York.

Kruger, D.H. and Bickle, T.A. (1983) Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol Rev* **47**: 345-60

Kuo, M.Y., Yang, M.K., Chen, W.P. and Kuo, T.T. (2000) High-frequency interconversion of turbid and clear plaque strains of bacteriophage f1 and associated host cell death. *Can J Microbiol* **46**: 841-7

Kwan, T., Liu, J., DuBow, M., Gros, P. and Pelletier, J. (2005) The complete genomes and proteomes of 27 *Staphylococcus aureus* bacteriophages. *Proc Natl Acad Sci U S A* **102**: 5174-9

Laird, W. and Groman, N. (1976a) Orientation of the tox gene in the prophage of corynebacteriophage beta. *J Virol* **19**: 228-31

Laird, W. and Groman, N. (1976b) Prophage map of converting corynebacteriophage beta. *J Virol* **19**: 208-19

Lawrence, J.G. and Roth, J.R. (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**: 1843-60

Lax, A.J., Chanter, N., Pullinger, G.D., Higgins, T., Staddon, J.M. and Rozengurt, E. (1990) Sequence analysis of the potent mitogenic toxin of *Pasteurella multocida*. *FEBS Lett* **277**: 59-64

Lazar, S. and Waldor, M.K. (1998) ToxR-independent expression of cholera toxin from the replicative form of CTXphi. *Infect Immun* **66**: 394-7

Leakey, A.K., Ulett, G.C. and Hirst, R.G. (1998) BALB/c and C57Bl/6 mice infected with virulent *Burkholderia pseudomallei* provide contrasting animal models for the acute and chronic forms of human melioidosis. *Microb Pathog* **24**: 269-75

Leclerc, H. and Sureau, P. (1956) Recherche des bacteriophages anti-bacille de whitmore dans les eaux stagnantes a Hanoi. *Bull. Soc. Path. Ex.* **56**: 874-882

Leelarasamee, A. (1998) *Burkholderia pseudomallei*: the unbeatable foe? *Southeast Asian J Trop Med Public Health* **29**: 410-5

Leelarasamee, A. and Bovornkitti, S. (1989) Melioidosis: review and update. *Rev Infect Dis* **11**: 413-25

Lehman, I.R. (1974) DNA ligase: structure, mechanism, and function. *Science* **186**: 790-7

Leiman, P.G., Chipman, P.R., Kostyuchenko, V.A., Mesyanzhinov, V.V. and Rossmann, M.G. (2004) Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118**: 419-29

Levin, B.R. and Tauxe, R.V. (1996) Cholera: nice bacteria and bad viruses. *Curr Biol* **6**: 1389-91

Lew, A.E. and Desmarchelier, P.M. (1993) Molecular typing of *Pseudomonas pseudomallei*: restriction fragment length polymorphisms of rRNA genes. *J Clin Microbiol* **31**: 533-9

Lieb, M. (1953) The establishment of lysogenicity in *Escherichia coli*. *J Bacteriol* **65**: 642-51

Lobet, Y., Cluff, C.W. and Cieplak, W., Jr. (1991) Effect of site-directed mutagenic alterations on ADP-ribosyltransferase activity of the A subunit of *Escherichia coli* heat-labile enterotoxin. *Infect Immun* **59**: 2870-9

Locht, C., Capiau, C. and Feron, C. (1989) Identification of amino acid residues essential for the enzymatic activities of pertussis toxin. *Proc Natl Acad Sci U S A* **86**: 3075-9

Loessner, M.J. (2005) Bacteriophage endolysins - current state of research and applications. *Curr Opin Microbiol* **8**: 480-7

Loessner, M.J., Goeppl, S. and Busse, M. (1991) Comparative inducability of bacteriophage in naturally lysogenesis and lysogenised strains of *Listeria* spp. by u.v. light and Mitomycin C. *Letters in Applied Microbiology* **12**: 196-199

Lu, Z., Breidt, F., Plengvidhya, V. and Fleming, H.P. (2003) Bacteriophage ecology in commercial sauerkraut fermentations. *Appl Environ Microbiol* **69**: 3192-202

Lunde, M., Aastveit, A.H., Blatny, J.M. and Nes, I.F. (2005) Effects of diverse environmental conditions on {phi}LC3 prophage stability in *Lactococcus lactis*. *Appl Environ Microbiol* **71**: 721-7

Luria, S.E., Delbruck, M. and Anderson, T.F. (1943) Electron Microscope Studies of Bacterial Viruses. *J Bacteriol* **46**: 57-77

Mackowiak, P.A. and Smith, J.W. (1978) Septicemic melioidosis. Occurrence following acute influenza A six years after exposure in Vietnam. *Jama* **240**: 764-6

Madera, C., Monjardin, C. and Suarez, J.E. (2004) Milk contamination and resistance to processing conditions determine the fate of *Lactococcus lactis* bacteriophages in dairies. *Appl Environ Microbiol* **70**: 7365-71

Madsen, S.M., Mills, D., Djordjevic, G., Israelsen, H. and Klaenhammer, T.R. (2001) Analysis of the genetic switch and replication region of a P335-type bacteriophage with an obligate lytic lifestyle on *Lactococcus lactis*. *Appl Environ Microbiol* **67**: 1128-39

Maniloff, J. and Ackermann, H.W. (1998) Taxonomy of bacterial viruses: establishment of tailed virus genera and the Order Caudovirales. *Arch Virol* **143**: 2051-63

Manna, D., Breier, A.M. and Higgins, N.P. (2004) Microarray analysis of transposition targets in *Escherichia coli*: the impact of transcription. *Proc Natl Acad Sci U S A* **101**: 9780-5

Manzeniuk, O., Volozhantsev, N.V. and Svetoch, E.A. (1994) Identification of the bacterium *Pseudomonas mallei* using *Pseudomonas pseudomallei* bacteriophages. *Mikrobiologiia* **63**: 537-44

Masui, S., Kamoda, S., Sasaki, T. and Ishikawa, H. (2000) Distribution and evolution of bacteriophage WO in *Wolbachia*, the endosymbiont causing sexual alterations in arthropods. *J Mol Evol* **51**: 491-7

Mathur, M.D., Vidhani, S. and Mehndiratta, P.L. (2003) Bacteriophage therapy: an alternative to conventional antibiotics. *J Assoc Physicians India* **51**: 593-6

Matsuda, T., Freeman, T.A., Hilbert, D.W., Duff, M., Fuortes, M., Stapleton, P.P. and Daly, J.M. (2005) Lysis-deficient bacteriophage therapy decreases endotoxin and inflammatory mediator release and improves survival in a murine peritonitis model. *Surgery* **137**: 639-46

McGrath, S., Fitzgerald, G.F. and van Sinderen, D. (2002) Identification and characterization of phage-resistance genes in temperate lactococcal bacteriophages. *Mol Microbiol* **43**: 509-20

Mecsas, J.J. and Strauss, E.J. (1996) Molecular mechanisms of bacterial virulence: type III secretion and pathogenicity islands. *Emerg Infect Dis* **2**: 270-88

Mehta, P., Casjens, S. and Krishnaswamy, S. (2004) Analysis of the lambdoid prophage element e14 in the *E. coli* K-12 genome. *BMC Microbiol* **4**: 4

Mel, S.F. and Mekalanos, J.J. (1996) Modulation of horizontal gene transfer in pathogenic bacteria by in vivo signals. *Cell* **87**: 795-8

Melton, A.R. and Weiss, A.A. (1989) Environmental regulation of expression of virulence determinants in *Bordetella pertussis*. *J Bacteriol* **171**: 6206-12

Merril, C.R., Biswas, B., Carlton, R., Jensen, N.C., Creed, G.J., Zullo, S. and Adhya, S. (1996) Long-circulating bacteriophage as antibacterial agents. *Proc Natl Acad Sci U S A* **93**: 3188-92

Miao, E.A. and Miller, S.I. (1999) Bacteriophages in the evolution of pathogen-host interactions. *Proc Natl Acad Sci U S A* **96**: 9452-4

Michiels, T. and Cornelis, G.R. (1991) Secretion of hybrid proteins by the *Yersinia* Yop export system. *J Bacteriol* **173**: 1677-85

Miles, A.A., Misra, S.S. and Irwin, J.O. (1938) The estimation of the bactericidal power of the blood. *J. Hyg., Camb.,* **38**: 732

Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T. and Ruger, W. (2003) Bacteriophage T4 genome. *Microbiol Mol Biol Rev* **67**: 86-156, table of contents

Miller, R.V. (2001) Environmental bacteriophage-host interactions: factors contribution to natural transduction. *Antonie Van Leeuwenhoek* **79**: 141-7

Miller, R.V. and Ripp, S. (1998) The importance of pseudolysogeny to *in situ* bacteriophage-host interactions. In: M. Syvanen and C.I. Kado (Eds), Horizontal gene transfer, pp. 179-191. Chapman and Hall, London.

Mirold, S., Rabsch, W., Rohde, M., Stender, S., Tschape, H., Russmann, H., Igwe, E. and Hardt, W.D. (1999) Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella typhimurium* strain. *Proc Natl Acad Sci U S A* **96**: 9845-50

Mirold, S., Rabsch, W., Tschape, H. and Hardt, W.D. (2001) Transfer of the Salmonella type III effector sopE between unrelated phage families. *J Mol Biol* **312**: 7-16

Mohamed, R., Nathan, S., Embi, N., Razak, N. and Ismail, G. (1989a) Inhibition of macromolecular synthesis in cultured macrophages by *Pseudomonas pseudomallei* exotoxin. *Microbiol Immunol* **33**: 811-20

Mohamed, R., Periasamy, M., Ismail, G. and Embi, N. (1989b) An ADP-ribosylating activity of the *Pseudomonas pseudomallei* exotoxin. *Conference Proceedings* **C57**

Morita, M., Tanji, Y., Orito, Y., Mizoguchi, K., Soejima, A. and Unno, H. (2001) Functional analysis of antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin against Gram negative bacteria. *FEBS Lett* **500**: 56-9

Morrow, J.F. (1979) Recombinant DNA techniques. *Methods Enzymol* **68**: 3-24

Munro, J., Oakey, J., Bromage, E. and Owens, L. (2003) Experimental bacteriophage-mediated virulence in strains of *Vibrio harveyi*. *Dis Aquat Organ* **54**: 187-94

Nathans, D. and Smith, H.O. (1975) Restriction endonucleases in the analysis and restructuring of dna molecules. *Annu Rev Biochem* **44**: 273-93

Nau, R. and Eiffert, H. (2002) Modulation of release of proinflammatory bacterial compounds by antibacterials: potential impact on course of inflammation and outcome in sepsis and meningitis. *Clin Microbiol Rev* **15**: 95-110

Nelson, D. (2004) Phage taxonomy: we agree to disagree. *J Bacteriol* **186**: 7029-31

Nelson, D., Loomis, L. and Fischetti, V.A. (2001) Prevention and elimination of upper respiratory colonization of mice by Group A streptococci by using a bacteriophage lytic enzyme. *Proc Natl Acad Sci U S A* **98**: 4107-12

Newland, J.W., Strockbine, N.A., Miller, S.F., O'Brien, A.D. and Holmes, R.K. (1985) Cloning of Shiga-like toxin structural genes from a toxin converting phage of *Escherichia coli*. *Science* **230**: 179-81

Ngauy, V., Lemeshev, Y., Sadkowski, L. and Crawford, G. (2005) Cutaneous melioidosis in a man who was taken as a prisoner of war by the Japanese during World War II. *J Clin Microbiol* **43**: 970-2

Nikaido, H. and Rosenberg, E.Y. (1981) Effect on solute size on diffusion rates through the transmembrane pores of the outer membrane of *Escherichia coli*. *J Gen Physiol* **77**: 121-35

Nikaido, H., Rosenberg, E.Y. and Foulds, J. (1983) Porin channels in *Escherichia coli*: studies with beta-lactams in intact cells. *J Bacteriol* **153**: 232-40

Normark, S., Nilsson, C. and Normark, B.H. (2005) Microbiology. A pathogen attacks while keeping up defense. *Science* **307**: 1211-2

Nutt, D. and Lingford-Hughes, A. (2004) Infecting the brain to stop addiction? *Proc Natl Acad Sci U S A* **101**: 11193-4

Oakey, H.J., Cullen, B.R. and Owens, L. (2002) The complete nucleotide sequence of the *Vibrio harveyi* bacteriophage VHML. *J Appl Microbiol* **93**: 1089-98

Oakey, H.J., Cullen, B.R. and Owens, L. (2005) A hypothetical model for VHML phage conversion of *Vibrio harveyi*. In: P. Walker, R. Lester and M.G. Bondad-Reantaso (Eds), Diseases in Asian Aquaculture V, pp. 457-464.

Oakey, H.J. and Owens, L. (2000) A new bacteriophage, VHML, isolated from a toxin-producing strain of *Vibrio harveyi* in tropical Australia. *J Appl Microbiol* **89**: 702-9

Ochs, H.D., Davis, S.D. and Wedgwood, R.J. (1971) Immunologic responses to bacteriophage phi-X 174 in immunodeficiency diseases. *J Clin Invest* **50**: 2559-68

O'Flynn, G., Ross, R.P., Fitzgerald, G.F. and Coffey, A. (2004) Evaluation of a cocktail of three bacteriophages for biocontrol of *Escherichia coli* O157:H7. *Appl Environ Microbiol* **70**: 3417-24

Ogunseitan, O.A., Sayler, G.S. and Miller, R.V. (1992) Application of DNA probes to analysis of bacteriophage distribution patterns in the environment. *Appl Environ Microbiol* **58**: 2046-52

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K. and Sugino, A. (1968) Mechanism of DNA chain growth. I. Possible discontinuity and unusual secondary structure of newly synthesized chains. *Proc Natl Acad Sci U S A* **59**: 598-605

Oki, M., Kakikawa, M., Yamada, K., Taketo, A. and Kodaira, K.I. (1996) Cloning, sequence analysis, and expression of the genes encoding lytic functions of Bacteriophage phi g1e. *Gene* **176**: 215-23

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L. and Adhya, S. (2005) Switches in bacteriophage lambda development. *Annu Rev Genet* **39**: 409-29

Osborn, M.J. and Wu, H.C. (1980) Proteins of the outer membrane of Gram negative bacteria. *Annu Rev Microbiol* **34**: 369-422

Paisano, A.F., Spira, B., Cai, S. and Bombana, A.C. (2004) In vitro antimicrobial effect of bacteriophages on human dentin infected with *Enterococcus faecalis* ATCC 29212. *Oral Microbiol Immunol* **19**: 327-30

Parfitt, T. (2005) Georgia: an unlikely stronghold for bacteriophage therapy. *Lancet* **365**: 2166-7

Paul, J.H., Sullivan, M.B., Segall, A.M. and Rohwer, F. (2002) Marine phage genomics. *Comp Biochem Physiol B Biochem Mol Biol* **133**: 463-76

Peltroche-Llacsahuanga, H. and Haase, G. (1999) A deadly thorn: a case of imported melioidosis. *Lancet* **353**: 1016-7

Pheiffer, B.H. and Zimmerman, S.B. (1983) Polymer-stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribooligonucleotides by T4 DNA ligase in polymer solutions. *Nucleic Acids Res* **11**: 7853-71

Popoff, M., Boquet, P., Gill, D.M. and Eklund, M.W. (1990) DNA sequence of exoenzyme C3, an ADP-ribosyltransferase encoded by *Clostridium botulinum* C and D phages. *Nucleic Acids Res* **18**: 1291

Pruzzo, C., Debbia, E.A. and Satta, G. (1980) Identification of the major adherence ligand of *Klebsiella pneumoniae* in the receptor for coliphage T7 and alteration of *Klebsiella* adherence properties by lysogenic conversion. *Infect Immun* **30**: 562-71

Ptashne, M. and Gilbert, W. (1970) Genetic repressors. *Sci Am* **222**: 36-44

Ptashne, M., Johnson, A.D. and Pabo, C.O. (1982) A genetic switch in a bacterial virus. *Sci Am* **247**: 128-30, 132, 134-40

Puthucheary, S.D., Vadivelu, J., Ce-Cile, C., Kum-Thong, W. and Ismail, G. (1996) Short report: Electron microscopic demonstration of extracellular structure of Burkholderia pseudomallei. *Am J Trop Med Hyg* **54**: 313-4

Raina, J.L. (1981) Purification of *Streptococcus* group C bacteriophage lysin. *J Bacteriol* **145**: 661-3

Randerson, J. (2003) Virus beats food bug. *New Scientist* **178**(2392): 22

Raina, J.L. (1981) Purification of *Streptococcus* group C bacteriophage lysin. *J Bacteriol* **145**: 661-3

Ratnaningsih, E., Dharmsthiti, S., Krishnapillai, V., Morgan, A., Sinclair, M. and Holloway, B.W. (1990) A combined physical and genetic map of *Pseudomonas aeruginosa* PAO. *J Gen Microbiol* **136 (Pt 12)**: 2351-7

Rebar, E.J. and Pabo, C.O. (1994) Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263**: 671-3

Reed, L.J. and Muench, H. (1938) A simple method of estimating fifty per cent endpoints. *The american Journal of Hygeine* **27**: 493-97

Rietsch, A., Wolfgang, M.C. and Mekalanos, J.J. (2004) Effect of metabolic imbalance on expression of type III secretion genes in *Pseudomonas aeruginosa*. *Infect Immun* **72**: 1383-90

Rimington, R.A. (1962) Melioidosis in north Queensland. *Med J Aust* **49(1)**: 50-3

Rohwer, F. (2003) Global phage diversity. *Cell* **113**: 141

Rohwer, F. and Edwards, R. (2002) The Phage Proteomic Tree: a genome-based taxonomy for phage. *J Bacteriol* **184**: 4529-35

Rosner, J.L. (1972) Formation, induction, and curing of bacteriophage P1 lysogens. *Virology* **48**: 679-89

Rubin, E.J., Lin, W., Mekalanos, J.J. and Waldor, M.K. (1998) Replication and integration of a *Vibrio cholerae* cryptic plasmid linked to the CTX prophage. *Mol Microbiol* **28**: 1247-54

Rusckowski, M., Gupta, S., Liu, G., Dou, S. and Hnatowich, D.J. (2004) Investigations of a (99m)Tc-labeled bacteriophage as a potential infection-specific imaging agent. *J Nucl Med* **45**: 1201-8

Russel, M., Lowman, H.B. and Clackson, T. (2004) Introduction to phage biology and phage display. In: T. Clackson and H.B. Lowman (Eds), Phage display. Vol. 266. Oxford university press, Oxford.

Sambrook, J. and Russell, D.W. (2001) Molecular Cloning; A Laboratory Manual. 3 ed, edited by J. Sambrook and D.W. Russell. 3 vols. Cold Spring Harbour Laboratory Press, Cold Spring Harbor.

Sandmeier, H. (1994) Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibres. *Mol Microbiol* **12**: 343-50

Scandella, D. and Arber, W. (1974) An *Escherichia coli* mutant which inhibits the injection of phage lambda DNA. *Virology* **58**: 504-13

Schierova, M., Linka, M. and Pazoutova, S. (2000) Sulfate assimilation in *Aspergillus terreus*: analysis of genes encoding ATP-sulfurylase and PAPS-reductase. *Curr Genet* **38**: 126-31

Schuch, R., Nelson, D. and Fischetti, V.A. (2002) A bacteriolytic agent that detects and kills *Bacillus anthracis*. *Nature* **418**: 884-9

Scotland, S.M., Smith, H.R., Willshaw, G.A. and Rowe, B. (1983) Vero cytotoxin production in strain of *Escherichia coli* is determined by genes carried on bacteriophage. *Lancet* **2**: 216

Seed, K.D. and Dennis, J.J. (2005) Isolation and characterization of bacteriophages of the *Burkholderia cepacia* complex. *FEMS Microbiol Lett* **251**: 273-80

Shearman, C.A., Jury, K.L. and Gasson, M.J. (1994) Controlled expression and structural organization of a *Lactococcus lactis* bacteriophage lysin encoded by two overlapping genes. *Appl Environ Microbiol* **60**: 3063-73

Shearman, C., Underwood, H., Jury, K. and Gasson, M. (1989) Cloning and DNA sequence analysis of a *Lactococcus* bacteriophage lysin gene. *Mol Gen Genet* **218**: 214-21

Sheehan, M.M., Stanley, E., Fitzgerald, G.F. and van Sinderen, D. (1999) Identification and characterization of a lysis module present in a large proportion of bacteriophages infecting *Streptococcus thermophilus*. *Appl Environ Microbiol* **65**: 569-77

Shrestha, N., Sharma, S., Khanal, B., Bhatta, N. and Dhakal, S. (2005) Melioidosis imported into Nepal. *Scand J Infect Dis* **37**: 64-6

Silhavey, T.J., Berman, M.L. and L.W., E. (1984) Experiments with Gene Fusion. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Sillankorva, S., Oliveira, R., Vieira, M.J., Sutherland, I.W. and Azeredo, J. (2004) Bacteriophage Phi S1 infection of *Pseudomonas fluorescens* planktonic cells versus biofilms. *Biofouling* **20**: 133-8

Slack, M.P. and Wheldon, D.B. (1978) A simple and safe volumetric alternative to the method of Miles, Misra and Irwin for counting viable bacteria. *J Med Microbiol* **11**: 541-5

Smith, C.J., Allen, J.C., Noor Embi, M., Othman, O., Razak, N. and Ismail, G. (1987a) Human melioidosis: an emerging medical problem. *MIRCEN Journal* **3**: 343-366

Smith, H.W., Green, P. and Parsell, Z. (1983) Vero cell toxins in *Escherichia coli* and related bacteria: transfer by phage and conjugation and toxic action in laboratory animals, chickens and pigs. *J Gen Microbiol* **129 (Pt 10)**: 3121-37

Smith, H.W., Huggins, M.B. and Shaw, K.M. (1987b) The control of experimental *Escherichia coli* diarrhoea in calves by means of bacteriophages. *J Gen Microbiol* **133**: 1111-26

Smith, P.B., Cherry, W.B. (1957) Identification of *malleomyces* by specific bacteriophages. **74**: 668-672

Snyder, L. and Champness, W. (1997) Bacteriophages. In: L. Snyder and W. Champness (Eds), Molecular genetics of bacteria. American Society for Microbiology, Washington DC.

Srinivasan, A., Kraus, C.N., DeShazer, D., Becker, P.M., Dick, J.D., Spacek, L., Bartlett, J.G., Byrne, W.R. and Thomas, D.L. (2001) Glanders in a military research microbiologist. *N Engl J Med* **345**: 256-8

Stocsits, R.R., Hofacker, I.L., Fried, C. and Stadler, P.F. (2005) Multiple sequence alignments of partially coding nucleic acid sequences. *BMC Bioinformatics* **6**: 160

Strauch, E., Lurz, R. and Beutin, L. (2001) Characterization of a Shiga toxin-encoding temperate bacteriophage of *Shigella sonnei*. *Infect Immun* **69**: 7588-95
Strauch, E., Schaudinn, C. and Beutin, L. (2004) First-time isolation and characterization of a bacteriophage encoding the Shiga toxin 2c variant, which is globally spread in strains of *Escherichia coli* O157. *Infect Immun* **72**: 7030-9
Streisinger, G., Emrich, J. and Stahl, M.M. (1967) Chromosome structure in phage T4, III. Terminal redundancy and length determination. *Proc Natl Acad Sci U S A* **57**: 292-295
Strockbine, N.A., Marques, L.R., Newland, J.W., Smith, H.W., Holmes, R.K. and O'Brien, A.D. (1986) Two toxin-converting phages from *Escherichia coli* O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. *Infect Immun* **53**: 135-40
Su, M.T., Venkatesh, T.V. and Bodmer, R. (1998) Large- and small-scale preparation of bacteriophage lambda lysate and DNA. *Biotechniques* **25**: 44-6
Sulakvelidze, A. (2005) Phage therapy: an attractive option for dealing with antibiotic-resistant bacterial infections. *Drug Discov Today* **10**: 807-9
Summer, E.J., Gonzalez, C.F., Carlisle, T., Mebane, L.M., Cass, A.M., Savva, C.G., LiPuma, J. and Young, R. (2004) *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *J Mol Biol* **340**: 49-65
Suttle, C.A. (1994) The significance of viruses to mortality in aquatic microbial communities. *Microb Ecol* **28**: 237-243
Takada, T., Iida, K. and Moss, J. (1995) Conservation of a common motif in enzymes catalyzing ADP-ribose transfer. Identification of domains in mammalian transferases. *J Biol Chem* **270**: 541-4
Tang, S.L., Nuttall, S., Ngui, K., Fisher, C., Lopez, P. and Dyall-Smith, M. (2002) HF2: a double-stranded DNA tailed haloarchaeal virus with a mosaic genome. *Mol Microbiol* **44**: 283-96
Taylor, A.L. (1963) Bacteriophage-Induced Mutation in *Escherichia coli*. *Proc Natl Acad Sci U S A* **50**: 1043-51
Thirion, J.P. and Hofnung, M. (1972) On some genetic aspects of phage lambda resistance in *E. coli* K12. *Genetics* **71**: 207-16
Tiemann, B., Depping, R., Gineikiene, E., Kaliniene, L., Nivinskas, R. and Ruger, W. (2004) ModA and ModB, two ADP-ribosyltransferases encoded by bacteriophage T4: catalytic properties and mutation analysis. *J Bacteriol* **186**: 7262-72
Trepel, J.B., Chuang, D.M. and Neff, N.H. (1977) Transfer of ADP-ribose from NAD to choleragen: a subunit acts as catalyst and acceptor protein. *Proc Natl Acad Sci U S A* **74**: 5440-2
Tsujimoto, H., Gotoh, N., Yamagishi, J., Oyamada, Y. and Nishino, T. (1997) Cloning and expression of the major porin protein gene opcP of *Burkholderia* (formerly *Pseudomonas*) *cepacia* in *Escherichia coli*. *Gene* **186**: 113-8
Twort, F.W. (1915) An investigation on the nature of ultra-microscopic viruses. *The Lancet* **186**: 1241-1243
Uchida, T., Gill, D.M. and Pappenheimer, A.M., Jr. (1971) Mutation in the structural gene for diphtheria toxin carried by temperate phage. *Nat New Biol* **233**: 8-11

Ueda, K. and Komano, T. (1984) Sequence-specific DNA damage induced by reduced mitomycin C and 7-N-(p-hydroxyphenyl)mitomycin C. *Nucleic Acids Res* **12**: 6673-83

Uetz, P., Rajagopala, S.V., Dong, Y.A. and Haas, J. (2004) From ORFeomes to protein interaction maps in viruses. *Genome Res* **14**: 2029-33

Ulett, G.C. (1996) The pathogenesis and treatment of melioidosis in relatively resistant c57BL/6 and highly susceptable BALB/c mice. Thesis, B. Biomed Sc Hons., Microbiology and Immunology, James Cook University.

Ulett, G.C., Currie, B.J., Clair, T.W., Mayo, M., Ketheesan, N., Labrooy, J., Gal, D., Norton, R., Smith, C.A., Barnes, J., Warner, J. and Hirst, R.G. (2001) *Burkholderia pseudomallei* virulence: definition, stability and association with clonality. *Microbes Infect* **3**: 621-31

Ulett, G.C., Ketheesan, N., Clair, T., Roberts, B., La Brooy, J., Norton, R., Ashhurst-Smith, C., Warner, J. and Hirst, R.G. (2000a) Cellular and Cytokine Responses in Experimental Melioidosis: Virulence Influences the Induction of TNF-a, IL-1B and IL-6 During the Innate Phase of Immunity. *2nd Congress of the Federation of Immunological Societies of Asia-Oceania*: 137-144

Ulett, G.C., Ketheesan, N. and Hirst, R.G. (2000b) Cytokine gene expression in innately susceptible BALB/c mice and relatively resistant C57BL/6 mice during infection with virulent *Burkholderia pseudomallei*. *Infect Immun* **68**: 2034-42

Van Dessel, W., Van Mellaert, L., Liesegang, H., Raasch, C., De Keersmaeker, S., Geukens, N., Lammertyn, E., Streit, W. and Anne, J. (2005) Complete genomic nucleotide sequence and analysis of the temperate bacteriophage VWB. *Virology* **331**: 325-37

Vegge, C.S., Brondsted, L., Neve, H., Mc Grath, S., van Sinderen, D. and Vogensen, F.K. (2005) Structural characterization and assembly of the distal tail structure of the temperate lactococcal bacteriophage TP901-1. *J Bacteriol* **187**: 4187-97

Verthe, K., Possemiers, S., Boon, N., Vaneechoutte, M. and Verstraete, W. (2004) Stability and activity of an *Enterobacter aerogenes*-specific bacteriophage under simulated gastro-intestinal conditions. *Appl Microbiol Biotechnol* **65**: 465-72

Vilagines, P., Sarrette, B. and Vilagines, R. (1982) Preformed magnesium hydroxide precipitate for second-step concentration of enteroviruses from drinking and surface waters. *Can J Microbiol* **28**: 783-7

Voelker, L.L. and Dybvig, K. (1998) Characterization of the lysogenic bacteriophage MAV1 from *Mycoplasma arthritidis*. *J Bacteriol* **180**: 5928-31

Wagner, P.L., Acheson, D.W. and Waldor, M.K. (1999) Isogenic lysogens of diverse shiga toxin 2-encoding bacteriophages produce markedly different amounts of shiga toxin. *Infect Immun* **67**: 6710-4

Wagner, P.L., Acheson, D.W. and Waldor, M.K. (2001) Human neutrophils and their products induce Shiga toxin production by enterohemorrhagic *Escherichia coli*. *Infect Immun* **69**: 1934-7

Waldor, M.K. and Mekalanos, J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910-4

Walkerpeach, C.R. and Pasloske, B.L. (2004) DNA bacteriophage as controls for clinical viral testing. *Clin Chem* **50**: 1970-1

Wang, I.N., Smith, D.L. and Young, R. (2000) Holins: the protein clocks of bacteriophage infections. *Annu Rev Microbiol* **54**: 799-825

Webb, J.S., Lau, M. and Kjelleberg, S. (2004) Bacteriophage and phenotypic variation in *Pseudomonas aeruginosa* biofilm development. *J Bacteriol* **186**: 8066-73

Weeks, C.R. and Ferretti, J.J. (1986) Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin) gene from *Streptococcus pyogenes* bacteriophage T12. *Infect Immun* **52**: 144-50

Weinbauer, M.G. and Rassoulzadegan, F. (2004) Are viruses driving microbial diversification and diversity? *Environ Microbiol* **6**: 1-11

Weinbauer, M.G. and Suttle, C.A. (1996) Potential significance of lysogeny to bacteriophage production and bacterial mortality in coastal water of the gulf of Mexico. *Appl Environ Microbiol* **62**: 4374-4380

Weitz, J.S., Hartman, H. and Levin, S.A. (2005) Coevolutionary arms races between bacteria and bacteriophage. *Proc Natl Acad Sci U S A* **102**: 9535-40

White, N.J. (2003) Melioidosis. *Lancet* **361**: 1715-22

Whitmore, A. (1913) An account of a Galnders-like disease occuring in Rangoon. *The Journal of Hygiene* **13**: 1-35

Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Orskov, F., Orskov, I. and Wilson, R.A. (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect Immun* **61**: 1619-29

Wiersinga, W.J., van der Poll, T., White, N.J., Day, N.P. and Peacock, S.J. (2006) Melioidosis: insights into the pathogenicity of *Burkholderia pseudomallei*. *Nat Rev Microbiol* **4**: 272-82

Wilkens, K., Tiemann, B., Bazan, F. and Ruger, W. (1997) ADP-ribosylation and early transcription regulation by bacteriophage T4. *Adv Exp Med Biol* **419**: 71-82

Williams, B.J., Golomb, M., Phillips, T., Brownlee, J., Olson, M.V. and Smith, A.L. (2002a) Bacteriophage HP2 of *Haemophilus influenzae*. *J Bacteriol* **184**: 6893-905

Williams, S.J., Senaratne, R.H., Mougous, J.D., Riley, L.W. and Bertozzi, C.R. (2002b) 5'-adenosinephosphosulfate lies at a metabolic branch point in mycobacteria. *J Biol Chem* **277**: 32606-15

Winstanley, C., Hales, B.A. and Hart, C.A. (1999) Evidence for the presence in *Burkholderia pseudomallei* of a type III secretion system-associated gene cluster. *J Med Microbiol* **48**: 649-56

Winstanley, C. and Hart, C.A. (2000) Presence of type III secretion genes in *Burkholderia pseudomallei* correlates with Ara(-) phenotypes. *J Clin Microbiol* **38**: 883-5

Wommack, K.E. and Colwell, R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* **64**: 69-114

Woods, D.E., Jeddeloh, J.A., Fritz, D.L. and DeShazer, D. (2002) *Burkholderia thailandensis* E125 harbors a temperate bacteriophage specific for *Burkholderia mallei*. *J Bacteriol* **184**: 4003-17

Wozniak, D.J., Hsu, L.Y. and Galloway, D.R. (1988) His-426 of the *Pseudomonas aeruginosa* exotoxin A is required for ADP-ribosylation of elongation factor II. *Proc Natl Acad Sci U S A* **85**: 8880-4

Yahr, T.L., Goranson, J. and Frank, D.W. (1996) Exoenzyme S of *Pseudomonas aeruginosa* is secreted by a type III pathway. *Mol Microbiol* **22**: 991-1003

Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L. and Treiber, G. (1970) Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large-scale virus purification. *Virology* **40**: 734-44

Yarmolinsky, M.B. (2004) Bacteriophage P1 in retrospect and in prospect. *J Bacteriol* **186**: 7025-8

Yee, K.C., Lee, M.K., Chua, C.T. and Puthucheary, S.D. (1988) Melioidosis, the great mimicker: a report of 10 cases from Malaysia. *J Trop Med Hyg* **91**: 249-54

Yokota, S., Hayashi, T. and Matsumoto, H. (1994) Identification of the lipopolysaccharide core region as the receptor site for a cytotoxin-converting phage, phi CTX, of *Pseudomonas aeruginosa*. *J Bacteriol* **176**: 5262-9

Yoon, S.S., Kim, J.W., Breidt, F. and Fleming, H.P. (2001) Characterization of a lytic *Lactobacillus plantarum* bacteriophage and molecular cloning of a lysin gene in Escherichia coli. *Int J Food Microbiol* **65**: 63-74

Young, R. (1992) Bacteriophage lysis: mechanism and regulation. *Microbiol Rev* **56**: 430-81

Zhang, X., McDaniel, A.D., Wolf, L.E., Keusch, G.T., Waldor, M.K. and Acheson, D.W. (2000) Quinolone antibiotics induce Shiga toxin-encoding bacteriophages, toxin production, and death in mice. *J Infect Dis* **181**: 664-70

APPENDIX 1

AGARS, CULTURE MEDIA AND GENERAL REAGENTS

1.1 Ashdown agar

tryptone (Oxoid, England)	12g
glycerol (Univar, Ajax Finechem, Australia)	32ml
crystal violet (0.1% aqueous - 1.1.1) (AnalaR, BDH Chemicals, Australia)	4ml
neutral red (1% aqueous - 1.1.3) (Difco Laboratories, England)	4ml
agar Technical no. 3 (Oxoid, England)	12g
double distilled water	800ml

Combine ingredients and boil for 15 minutes. Autoclave at 121°C for 15 minutes, cool to 50°C and add 2.0ml of 5 mg/ml gentamycin sulphate (65.5% pure - 1.1.2). Pour into plates and cool.

1.1.1 crystal violet (0.1%)

crystal violet (AnalaR, BDH Chemicals, Australia)	0.1g
double distilled water	Make up to 100ml

Dissolve crystal violet in double distilled water and autoclave at 121°C for 15 minutes.

1.1.2 gentamycin sulphate (5mg/ml)

gentamycin sulphate (G3632, Sigma Chemicals, Australia)	0.05g
double distilled water	Make up to 10ml

Dissolve gentamycin sulphate in double distilled water and filter sterilise through a 0.22µm filter (Millex®-GS, Millipore, Ireland).

1.1.3 neutral red (1% aqueous)

neutral red (Difco Laboratories, England)	1g
double distilled water	Make up to 100ml

Dissolve neutral red in double distilled water and autoclave at 121°C for 15 minutes.

1.2 brain heart infusion broth (BHIB)

brain heart infusion broth	37g
(Acumedia, USA)	
double distilled water	Make up to 1000ml

Make up to 950ml with double distilled water, adjust pH to 7.4 and make up volume to 1000ml. Autoclave at 121°C for 15 minutes.

1.3 $CaCl_2$ (0.1M)

$CaCl_2$	11.1g
(Univar, Ajax Finechem, Australia)	
doubly distilled water	Make up to 1000ml

Dissolve $CaCl_2$ in 800ml doubly distilled water. Make up to 1000ml with doubly distilled water. Autoclave at 121°C for 15 minutes.

1.4 CTAB/NaCl solution

NaCl	4.1g
(Univar, Ajax Finechem, Australia)	
hexadecyltrimethyl ammonium bromide (CTAB)	10g
(Sigma Chemicals, Australia)	
doubly distilled water	Make up to 100ml

Dissolve NaCl in 80ml doubly distilled water. Add CTAB slowly with heating and stirring. Make up to 100ml with doubly distilled water.

1.5 EDTA (0.5M) (pH8.0)

EDTA	186.12g
(Univar, Ajax Finechem, Australia)	
double distilled water	Make up to 1000ml

Combine ingredients and make up to 800ml with double distilled water. Adjust pH to 8.0. Make up to 1000ml with double distilled water when EDTA has dissolved. Autoclave at 121°C for 15 minutes

1.6 Luria-Bertani (LB) medium

tryptone	10g
(Oxoid, England)	
yeast extract	5g
(Oxoid, England)	
NaCl	10g
(Univar, Ajax Finechem, Australia)	
double distilled water	Make up to 1000ml

Make up to 950ml with double distilled water, adjust pH to 7.0 and bring volume up to 1000ml. Autoclave at 121°C for 15 minutes. When using

media for protocols involving Promega cloning kits, only 5g of NaCl was added to the media as per kit instructions.

1.7 LB agar

LB medium (1.5)	100ml
agar Technical no. 3 (Oxoid, England)	1.5g

Combine ingredients and boil for 15 minutes. Autoclave at 121°C for 15 minutes, cool to 50°C and pour into plates

1.8 NaCl (5M)

NaCl (Univar, Ajax Finechem, Australia)	292.2g
doubly distilled water	Make up to 1000ml

Dissolve NaCl in 800ml doubly distilled water and make up to 1000ml with doubly distilled water. Autoclave at 121°C for 15 minutes.

1.9 phosphate buffered saline (PBS) (pH 7.4)

NaCl (Univar, Ajax Finechem, Australia)	8g
Na_2HPO_4 (Univar, Ajax Finechem, Australia)	1.44g
KH_2PO_4 (Sigma Chemicals, USA)	0.24g
KCl (Univar, Ajax Finechem, Australia)	0.2g
double distilled water	Make up to 1000ml

Make up to 950ml with double distilled water, adjust pH to 7.2 and make up volume to 1000ml. Autoclave at 121°C for 15 minutes

1.10 sheep blood agar (SBA)

blood agar base #2 (Oxoid, Unipath Ltd., England)	20g
doubly distilled water	500ml
sheep blood (JCU)	25 ml

Boil blood agar base in doubly distilled water until dissolved, then autoclave at 121°C for 15 minutes. Allow to cool to 55°C and add blood under sterile conditions. Mix, pour into plates and cool.

1.11 six by loading dye

glycerol (Univar, Ajax Finechem, Australia)	30ml
0.5M ETDA (1.5) (Univar, Ajax Finechem, Australia)	2 ml
bromophenol blue (Progen, Australia)	0.15g
orange G (Sigma Chemicals, Australia)	0.2g
xylene cyanole FF (Sigma Chemicals Australia)	0.075g
doubly distilled water	Make up to 100ml

Combine ingredients and make up to 100ml with double distilled water.

1.12 SOC buffer

Tryptone (Oxoid, England)	2.0g
yeast extract (Oxoid, England)	0.5g
1M NaCl (1.12.1)	1.0ml
1M KCl (1.12.2)	0.25ml
2M Mg^{2+} stock (1.12.3)	1.0ml
2M glucose (1.12.4)	1.0ml
double distilled water	Make up to 100ml

Combine first four ingredients and make up to 97ml with double distilled water (dissolve). Autoclave at 121°C for 15 minutes and cool to room temperature. Add 2M Mg^{2+} stock and 2M glucose and make up to 100ml with sterile double distilled water.

1.12.1 NaCl (1M)

NaCl (Univar, Ajax Finechem, Australia)	5.8g
doubly distilled water	Make up to 100ml

Dissolve NaCl in 80ml doubly distilled water and make up to 100ml with doubly distilled water. Autoclave at 121°C for 15 minutes.

1.12.2 KCl (1M)

KCl (Univar, Ajax Finechem, Australia)	7.5g
double distilled water	Make up to 100ml

Make up to 80ml with double distilled water and make up volume to 100ml. Autoclave at 121°C for 15 minutes

1.12.3 Mg^{2+} stock (2M)

$MgCl_2.6H_2O$	20.33g
(Univar, Ajax Finechem, Australia)	
$MgSO_4.7H_2O$	24.65g
(AnalaR, BDH Chemicals, Australia)	
double distilled water	Make up to 100ml

Combine ingredients and make up to 100ml with double distilled water. Filter sterilise.

1.12.4 glucose (2M)

glucose	36.0g
(AnalaR, BDH Chemicals, Australia)	
double distilled water	Make up to 100ml

Make up to 80ml with double distilled water and make up volume to 100ml. Filter sterilise.

1.13 SM buffer

NaCl	5.8g
(Univar, Ajax Finechem, Australia)	
$MgSO_4.7H_2O$	2.0g
(AnalaR, BDH Chemicals, Australia)	
1M tris-HCl (pH 7.5)(1.21)	50ml
2% gelatin (1.13.1)	5ml
double distilled water	Make up to 1000ml

Combine ingredients and make up to 1000ml with double distilled water. Aliquot as required and autoclave at 121°C for 15 minutes

1.13.1 2% gelatin

gelatin	0.2g
(Univar, Ajax Finechem, Australia)	
double distilled water	10ml

Dissolve gelatin in double distilled water with heating and stirring. Autoclave at 121°C for 15 minutes

1.14 SM buffer - modified (pH6.3, 0.4M KCl)

NaCl	5.8g
(Univar, Ajax Finechem, Australia)	
$MgSO_4.7H_2O$	2.0g
(AnalaR, BDH Chemicals, Australia)	
1M tris-HCl (pH 7.5)(1.21)	50ml
2% gelatin (1.13.1)	5ml
KCl	29.82g
(Univar, Ajax Finechem, Australia)	
double distilled water	Make up to 1000ml

Combine ingredients and make up to 800ml with double distilled water. Adjust pH to 6.3. Make up to 1000ml with double distilled water. Aliquot as required and autoclave at 121°C for 15 minutes

1.15 sodium dodecyl sulphate (SDS) (10% solution)

SDS	10g
(AnalaR, BDH Chemicals, Australia)	
doubly distilled water	Make up to 100ml

Dissolve SDS in 80ml doubly distilled water. Make up to 100ml with doubly distilled water.

1.16 sodium acetate (10% solution, pH5.2)

sodium acetate	10g
(AnalaR, BDH Chemicals, Australia)	
doubly distilled water	Make up to 100ml

Dissolve sodium acetate in 50ml doubly distilled water. Adjust pH to 5.2 with acetic acid. Make up to 100ml with doubly distilled water.

1.17 tris-acetate (TAE) buffer (50×)

tris	242g
(AnalaR, BDH Chemicals, Australia)	
glacial acetic acid	57.1ml
(Univar, Ajax Finechem, Australia)	
0.5M EDTA (pH8.0)(1.5)	100ml
double distilled water	Make up to 1000ml

Combine ingredients and make up to 1000ml with double distilled water. Autoclave at 121°C for 15 minutes

1.18 tris borate (TBE) buffer (5x)

tris	5.4g
(AnalaR, BDH Chemicals, Australia)	
boric acid	2.75 g
(Univar, Ajax Finechem, Australia)	
0.5M EDTA (pH8.0)(1.5)	2.0ml
double distilled water	Make up to 100ml

Combine ingredients and make up to 100ml with double distilled water. Autoclave at 121°C for 15 minutes. Dilute 1 in 10 (0.5x) for use.

1.19 tris-EDTA (TE) buffer

1M tris-HCl (1.21)	10ml
0.5M EDTA (1.5)	20ml
doubly distilled water	Make up to 1000ml

Combine ingredients and make up to 800ml with double distilled water. Adjust pH to 8.0. Make up to 1000ml with double distilled water. Autoclave at 121°C for 15 minutes.

1.20 tris-glucose buffer

1M tris-HCl (1.20)	50ml
glucose	50g
(Univar, Ajax Finechem, Australia)	
doubly distilled water	Make up to 1000ml

Combine ingredients and make up to 1000ml with double distilled water. Filter sterilise through a 0.22µm filter (Millex®-GS, Millipore, Ireland)

1.21 tris-HCl (1M) (pH 7.5)

tris	12.14g
(AnalaR, BDH Chemicals, Australia)	
double distilled water	100ml

Dissolve tris in 80ml double distilled water and adjust pH to 7.5. Make up to 100ml with double distilled water and autoclave at 121°C for 15 minutes.

1.22 tryptone soya broth (TSB)

tryptone soya broth	30g
(Oxoid, England)	
double distilled water	Make up to 1000ml

Make up to 950ml with double distilled water, adjust pH to 7.3 and make up volume to 1000ml. Autoclave at 121°C for 15 minutes.

APPENDIX 2

COLLATED PLAQUE FORMATION DATA

Table A2.1. Isolate name key for table A.2.2. All isolates are *B. pseudomallei* except for *B. thailandensis*. All isolates are in the JCU Microbiology bacterial collection.

Isolate name	Isolate Code
NCTC 13178	1
#2	2
#3	3
#4	4
#5	5
#6	6
#7	7
#8	8
#9	9
#10	10
#11	11
#12	12
#13	13
#14	14
#15	15
#16	16
#17	17
#18	18
#19	19
#20	20
#21	21
B. thailandensis	22
#23	23
#24	24
#25	25
#26	26
#27	27
#28	28
#29	29
#30	30
#31	31
#32	32
#33	33
#34	34
NCTC13179	35
ATCC 23343	36
NAFC	37
#38	38
#39	39
Ts5	40
Ts21	41
#73	42
#70	43
E1	44
E2	45
E3	46
E4	47
#69	48
#83	49
#82	50

Table A.2.2. Raw cross-spotting data. Isolate spotted on plate is in the form of broth supernatant. Scores are; 0-no interaction, 1-thinning, 2-complete clearing, 3-possible plaques, 4-definite plaques. The isolate name matching code for isolates 1-50 is described in table A.2.1. This table is described in four quarters due to size.

code of isolate plated	Code of isolate spotted on plate																								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
1	1	1	2	1	1	2	1	1	1	1	2	1	2	1	1	1	1	1	2	1	1	1	2	2	1
2	4	1	2	1	1	1	2	2	1	1	1	1	2	4	4	1	2	2	2	1	1	1	1	2	1
3	2	2	2	2	1	1	2	2	1	2	1	2	2	2	2	1	2	2	2	2	2	2	2	2	2
4	2	1	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	2	2	2	1	1	1	1	1
5	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	1	2	0	2	2	2
6	1	1	1	1	2	1	1	1	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	3
7	2	2	2	1	1	2	2	2	1	1	1	1	2	2	1	0	1	2	1	1	1	1	1	1	1
8	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
9	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1
10	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
12	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	0	0	0	1	1	1	1	1	1	2	1	1	1	1	1	1	1	0	1	2
15	1	1	2	1	1	0	0	1	1	1	1	1	1	2	2	1	1	0	0	0	1	1	0	0	1
16	1	1	1	1	1	1	1	1	2	2	1	1	1	2	4	2	1	2	1	1	2	1	1	1	1
17	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	0	0	1	0	1	1
18	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	0	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	0	0	0	0	0	4	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1
21	0	1	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
22	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	0	1	1	1	1	1	1	1
23	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1	1	1	1	1
24	1	1	1	1	0	0	3	0	0	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1
25	1	1	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Table A.2.2 continued

Code of isolate plated	Code of isolate spotted on plate																							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**
26	0	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	1	1	1	1	0	1	1	1	1
27	4	1	0	2	1	0	1	1	1	2	1	0	1	2	4	2	1	1	1	1	1	1	1	0	0
28	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	1	0	1	1	0	1
29	2	2	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	0	1	1	1	2
30	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
31	0	0	0	0	1	3	1	0	1	1	1	1	1	0	0	1	1	0	1	1	0	1	1	2	2
32	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	2
33	0	1	1	1	0	4	1	1	1	1	0	0	1	1	1	1	1	1	1	0	0	0	1	1	1
34	1	1	1	1	1	1	1	0	0	1	1	0	0	1	0	0	1	0	1	1	0	0	0	0	1
35	1	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	0	0	0	0	2	1	1	1
36	4	2	2	2	2	4	2	1	1	0	2	1	2	1	1	0	1	1	2	2	2	1	1	1	1
37	0	2	4	4	4	4	2	2	4	2	2	2	2	1	3	1	1	2	1	2	2	4	1	1	1
38	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1
39	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
40	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
41	1	2	1	1	0	1	2	1	1	2	1	1	1	1	1	4	1	1	1	3	1	3	1	1	1
42	1	1	2	1	1	1	2	1	2	1	1	2	0	2	1	0	1	1	1	1	0	0	1	1	1
43	1	1	1	1	1	1	1	2	1	1	1	1	0	1	0	1	1	1	1	0	0	0	1	1	0
44	0	0	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	3	1	1	3
45	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	1	1	1	0	0	1	1	1
46	1	1	0	0	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1
47	1	1	1	0	1	1	1	1	0	1	0	0	0	1	0	1	1	1	1	1	0	1	1	1	1
48	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
49	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1
50	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.2.2 continued

code of isolate plated	Code of isolate spotted on plate																								
	26	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1
2	1	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	2	0	0	0	1	0	1
3	2	0	0	1	0	0	1	0	0	1	0	2	0	0	1	2	2	0	0	0	0	0	0	1	1
4	1	0	0	1	0	0	0	0	0	4	1	0	1	1	0	1	1	0	0	0	0	1	0	0	0
5	2	1	1	1	0	1	2	1	2	1	1	1	1	1	2	2	0	1	1	1	1	1	1	1	1
6	1	1	1	1	0	1	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	2	1	0	1	2	2	2	1	1	1	2	2	1	1	0	1	1	1	2	1	1	1	1
8	0	1	2	1	0	1	2	2	2	1	1	2	1	2	2	2	2	1	1	1	2	2	1	0	0
9	0	2	1	2	0	1	4	4	4	4	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1
10	1	0	0	0	0	1	2	2	2	0	0	1	1	2	2	0	2	1	1	1	1	1	2	1	1
11	1	0	0	0	0	1	2	2	1	0	0	1	2	2	0	0	0	1	2	1	1	0	0	1	1
12	1	0	0	0	0	1	1	2	2	0	0	0	0	0	0	0	0	1	2	1	1	1	0	2	1
13	1	0	0	0	0	0	0	0	0	2	1	0	0	0	0	2	1	1	2	0	1	2	1	1	0
14	0	0	0	0	0	0	0	1	1	0	0	4	0	1	0	0	0	0	0	0	0	0	1	0	1
15	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0	0	0
16	1	0	4	0	0	0	1	2	2	0	0	1	1	2	1	0	0	1	2	2	1	0	0	0	0
17	1	1	1	1	0	1	2	2	2	0	0	1	1	0	1	0	0	0	0	1	1	1	0	0	0
18	1	1	2	1	0	0	1	2	0	2	0	0	1	1	1	0	0	0	0	1	0	0	0	0	0
19	1	4	4	0	0	2	2	2	2	1	2	2	2	2	2	2	0	1	2	2	2	2	2	1	1
20	1	0	1	0	0	3	0	3	3	4	2	0	0	0	0	0	0	0	0	0	0	0	2	0	0
21	1	1	1	1	1	1	2	2	1	1	1	2	2	2	1	1	2	1	1	1	1	1	1	1	1
22	1	1	0	0	0	0	2	0	0	0	0	2	2	0	0	1	1	0	0	2	2	1	0	0	0
23	1	0	0	0	0	0	2	2	1	1	0	0	2	2	2	2	1	0	0	1	2	1	1	0	0
24	1	1	1	1	0	0	2	2	2	1	1	0	2	1	2	1	1	1	1	1	1	1	1	1	1
25	1	0	0	0	0	1	1	2	1	1	0	1	2	2	1	2	1	1	2	1	1	2	1	0	0

Table A.2.2 continued

code of isolate plated	Code of isolate spotted on plate																							
	26	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
26	1	1	1	1	4	1	4	4	1	1	1	0	0	4	0	0	0	0	0	4	0	3	0	0	0
27	0	0	1	1	0	0	4	0	4	2	1	4	4	2	2	4	1	2	2	2	1	1	1	1	1
28	1	4	0	2	0	1	4	4	3	2	2	2	4	3	4	3	3	2	2	2	2	4	3	2	2
29	1	2	1	3	0	1	2	2	2	1	0	0	2	2	1	1	1	1	2	2	1	2	2	2	2
30	0	0	0	0	0	1	2	2	2	1	1	1	2	2	0	0	0	0	2	0	2	2	1	1	1
31	1	2	0	0	0	2	2	2	1	1	0	2	2	3	1	1	0	1	2	2	2	2	1	1	1
32	1	3	4	1	1	1	2	2	2	2	1	2	2	2	2	2	3	2	2	2	1	1	1	1	1
33	1	2	0	3	1	0	2	2	2	2	2	2	2	3	2	2	3	2	2	2	2	2	2	2	2
34	0	0	0	2	0	2	2	2	2	0	0	2	2	2	0	0	1	2	0	2	1	1	1	1	1
35	2	0	0	0	0	0	2	2	2	2	2	2	2	2	0	2	0	0	0	0	1	2	2	0	0
36	1	0	0	0	0	0	3	3	3	4	3	0	4	4	3	4	3	0	3	0	4	0	0	0	3
37	1	4	4	4	0	0	3	3	3	2	0	0	4	3	0	1	1	3	0	1	2	2	2	3	1
38	1	0	0	0	0	1	2	2	1	1	0	0	2	2	1	1	1	0	2	2	2	2	2	0	1
39	2	0	0	0	0	0	1	2	1	1	1	0	1	3	0	0	2	0	2	2	1	2	0	1	1
40	1	0	0	1	0	0	2	2	1	1	0	2	1	1	0	1	2	0	0	3	1	2	0	0	0
41	3	0	0	1	0	0	2	2	2	1	2	0	1	1	1	1	2	2	2	0	0	1	0	0	0
42	0	0	0	0	0	0	1	2	2	2	1	0	1	2	1	1	2	0	0	1	3	1	0	0	0
43	0	0	0	0	0	0	2	0	0	0	0	1	1	1	1	1	1	0	1	1	0	0	0	0	1
44	3	0	1	1	0	1	1	1	2	2	0	0	1	0	0	2	1	0	0	0	0	2	2	0	0
45	1	0	2	2	0	0	0	1	2	2	1	0	0	0	2	1	1	0	0	0	0	0	0	0	0
46	1	0	0	1	0	0	2	2	2	2	2	0	2	2	0	2	2	1	0	2	0	2	2	1	0
47	1	0	1	1	0	0	2	2	2	2	2	0	2	2	1	2	2	0	0	0	0	2	0	0	0
48	0	0	0	2	0	0	1	0	2	2	2	0	0	0	0	2	2	0	0	0	0	0	2	0	0
49	1	0	0	0	0	4	2	3	3	0	4	0	0	0	0	0	1	0	0	0	0	0	1	0	1
50	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0

Table A.2.3 Number of particular types of interactions each supernatant had with the 50 plated isolates from table A.2.2

	code of isolate spotted on plate															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
score 0	16	9	13	15	16	15	13	16	14	7	12	13	11	4	12	16
score 1	25	33	27	30	29	26	28	28	31	36	34	32	31	36	30	30
score 2	6	8	8	4	4	3	8	6	4	7	4	5	8	9	4	3
score 3	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0
score 4	3	0	2	1	1	5	0	0	1	0	0	0	0	1	3	1
score 2,3 or 4	9	8	10	5	5	9	9	6	5	7	4	5	8	10	8	4
score 2 or 4	9	8	10	5	5	8	8	6	5	7	4	5	8	10	7	4

	code of isolate spotted on plate															
	17	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
score 0	6	9	5	8	18	11	7	7	5	11	32	29	23	46	27	9
score 1	39	31	37	36	26	33	40	38	37	33	10	13	19	3	18	9
score 2	5	10	8	5	6	3	3	5	6	4	4	4	5	0	3	26
score 3	0	0	0	1	0	2	0	0	2	2	1	0	2	0	1	2
score 4	0	0	0	0	0	1	0	0	0	0	3	4	1	1	1	4
score 2,3 or 4	5	10	8	6	6	6	3	5	8	6	8	8	8	1	5	32
score 2 or 4	5	10	8	5	6	4	3	5	6	4	7	8	6	1	4	30

	code of isolate spotted on plate																	
	33	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
score 0	11	10	14	23	23	12	12	22	17	18	25	22	16	18	15	20	26	22
score 1	5	11	18	16	12	15	12	16	16	17	18	10	19	19	17	18	19	24
score 2	27	22	14	9	13	19	19	10	14	11	6	17	13	11	16	11	4	3
score 3	4	5	0	1	0	0	5	1	1	4	1	1	1	1	1	1	1	1
score 4	3	2	4	1	2	4	2	1	2	0	0	0	1	1	1	0	0	0
score 2,3 or 4	34	29	18	11	15	23	26	12	17	15	7	18	15	13	18	12	5	4
score 2 or 4	30	24	18	10	15	23	21	11	16	11	6	17	14	12	17	11	4	3

Table A.2.4 Number of particular types of interactions each plate had with the 50 supernatants from table A.2.2

	code of plated isolate															
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
score 0	19	16	15	16	3	1	3	15	7	7	14	15	14	23	30	12
score 1	24	22	11	26	20	42	31	24	36	33	31	28	32	24	16	25
score 2	7	9	24	7	27	6	16	11	3	10	5	7	4	2	4	11
score 3	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
score 4	0	3	0	1	0	0	0	0	4	0	0	0	0	1	0	2
score 2, 3 or 4	7	12	24	8	27	7	16	11	7	10	5	7	4	3	4	13
score 2 or 4	7	12	24	8	27	7	16	11	7	10	5	7	4	3	4	13

	code of plated isolate															
	17	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**
score 0	19	15	4	29	5	29	26	7	17	24	10	17	4	12	14	3
score 1	28	32	28	14	39	16	17	33	27	20	23	12	28	30	22	31
score 2	3	3	15	2	6	5	7	9	6	0	10	10	17	8	12	13
score 3	0	0	0	3	0	0	0	1	0	1	0	5	1	0	2	2
score 4	0	0	3	2	0	0	0	0	0	5	7	6	0	0	0	1
score 2, 3 or 4	3	3	18	7	6	5	7	10	6	6	17	21	18	8	14	16
score 2 or 4	3	3	18	7	6	5	7	10	6	6	17	21	18	8	14	16

	code of plated isolate																	
	33	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
score 0	9	20	19	13	7	11	11	11	11	16	21	18	32	17	18	39	16	47
score 1	19	20	18	12	12	30	32	32	25	23	27	24	13	21	22	4	28	3
score 2	18	10	13	10	14	9	6	5	10	10	2	5	5	12	10	7	2	0
score 3	3	0	0	8	7	0	1	1	3	1	0	3	0	0	0	0	2	0
score 4	1	0	0	7	10	0	0	1	1	0	0	0	0	0	0	0	2	0
score 2, 3 or 4	22	10	13	25	31	9	7	7	14	11	2	8	5	12	10	7	6	0
score 2 or 4	22	10	13	25	31	9	7	7	14	11	2	8	5	12	10	7	6	0

Table A.2.5 Examination of filter sterilised supernatants of *B. pseudomallei* isolates for the presence of bacteriophage by spotting on lawns of various *B. pseudomallei* and *B. thailandensis* bacteria. The presence of plaques is described by '+', no plaques is described by '-', possible plaques is described by '+?'. Isolates C4 and soil 36 do not have a code as they were not included in the cross-spotting assay. Greyed boxes are those interactions that were not examined .

code of isolate plated	code of isolate supernatant, filter sterilised and spotted on lawns, code described in table 1																														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	18	19	22	29	30	35	36	37	44	45	46	47	69	73	82	83	C4
1	-	-	-		+?	-	-	-	-	-	-	+?	-		-	+?	+	-	-	-		-	+	+			-	-	+?	-	-
4	+	-	-		-	-	-	-	-	+?	-	-	-		-	-	+?	+	+	-	-	+	+	+			-	+	-	-	-
6	+	-	-		-		-?	-	-	-	-	+?	-		+?	+	+	-	-	-	-?	+	+	+			+?	+	+	-	+
7	+	+	-		-	-		-	-	-	-	-	-		+	-	+	-?	-	-	-	+	+	+			+?	+	-	-	+
9	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	+	+	+	-	-	+	+	+		+	-	+	-	-	-
10		-	-	-	-			-	-	-	-	-	-	-	-	-	-	-	-	-		+	-	+		+	-	+	-	-	-
11		-	-		-			-	-	+	-	-	-		-	-	-	-	-	-		+?	-	-			-	-	-	-	+
12		-	-	-	-			-	-	-	-	-	-	-	-	-	+	-	-	-		+	-	+		+	-	+	-	-	-
13		-	-	-	-			-	-	-	-	-	-	-	-	-	+	-	-	+		+	+	+		+	-	+	+	-	+
14	-	-				-	-								-	-	-	-	-	-	-							-	-		
17	+	-		-		-	-							-	-	-	+	-	-	-	-				+			+	+		-
21	-	-		-		-	-							-	-	-	+	-	-	-	-				+			+	-		-
27	+	-		-		+	-							+	-	-	+	+	-	+	-				+			+	+		
28	-	-		-		+	-							-	-	-	-	-	-	+	-					-		-	+		-
29	+	-	-		-	-?	-	-	+	-	+	-	-		-	-	+	-	-	+	-	+?	-	+			-	+	+	+	-
41	+	-		-		-	-							-	-	-	+	-	-	-	-				+			+	+		-
73	-?	-	-		-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	+?	-	+			-	-	+	-	-
82	-	-	-		-	-?	-	-	-	+	+	-	-		-	-	-	-?	-	-	-	-					-	+	-	-	-
22	-	-	-		+?	-	-	?	+?	+	-	-	-		-	-	-	-	-	-	-	-	-	+			-	-	-	-	-
35	+	+	-		-	-	-	-	-	-	-	-	-		-	-	+	-	-	-	-	+	-	+			-	+	-	-	+?
soil 36	+	-		-		-	+							+	+	-	+	-	-	-	-				+			+	-		-
46	+?	-				-	+?								-	-	+	-	-	-	-							-	-		
47	+	-				-	-								-	-	+	-	-	-	-							-	-		
C4	-	-	-		-	-	+	-	-	-	-	-	-		+	+	+?	+?	+	-	-	-	-	-			-	-	+?	+?	-

APPENDIX 3

COMMERCIAL KIT PROTOCOLS ADAPTATIONS

3.1 Lambda Mini Kit Modified Method (Qiagen)

Steps from kit that are not included in my method;
step 1: propagation of bacteriophage, lysis with chloroform and centrifugation to remove bacterial debris.

Steps from kit that are modified in my method;
Step 2 is carried out using previously extracted bacteriophage, as per Section 4.2.2.1

All other steps carried out as per manufacturers instructions.

3.2 Nucleobond AX; (Machery Nagel)

Steps from kit that are not included in my method;
step 1: lysis of bacterial broth with chloroform and centrifugation to remove bacterial debris.

Steps from kit that are modified in my method;
step 2 and 3: This was either carried out on previously extracted and concentrated bacteriophage solution as per manufacturers instructions (Section 4.2.3), or eliminated if the concentrated bacteriophage was made up in modified SM buffer at pH 6.3 and containing KCl (Section 6.2.1.1).

All other steps carried out as per manufacturers instructions (Section 4.2.3) or a modification of steps 4-7 carried out as described (Section 6.2.1.1)

APPENDIX 4

MORBIDITY DATA FOR SELECTED *B. pseudomallei* ISOLATES AND ID_{50} CALCULATIONS

Table A.4.1 Morbidity data for BALB/c mice seven days after infection with isolate #13 (Section 5.3.1, assay F.1)

challenge (cfu/200µl)	dilution	morbidity ratio (#morbid/ #challenged)	accumulative morbid	accumulative not morbid	percent morbidity (%)
$1.2x10^8$	1:1 (neat)	not done			
$9.3x10^6$	0.049	38841	20	0	100
$8.7x10^5$	1:100	38841	15	0	100
$9.7x10^4$	1:1000	38841	10	0	100
$8.3x10^3$	1:10000	38840	5	1	80
$7.3x10^2$	1:100000	38837	1	5	20

Table A.4.2 Morbidity data for BALB/c mice seven days after infection with isolate NAFC (Section 5.3.1, assay A.1)

challenge (cfu/200µl	dilution	morbidity ratio (#morbid/ #challenged)	accumulative morbid	accumulative not morbid	percent morbidity (%)
$8.0x10^7$	1:1 (neat)	not done			
$7.4x10^5$	1:100	38841	10	0	100
$7.4x10^4$	1:1000	38841	5	0	100
$8.4x10^3$	1:10000	38837	0	5	0

Table A.4.3 Morbidity data for BALB/c mice seven days after infection with isolate NAFC + BupsΦ1 (Section 5.3.1, assay A.1)

challenge (cfu/200µl	dilution	morbidity ratio (#morbid/ #challenged)	accumulative morbid	accumulative not morbid	percent morbidity (%)
$5.6x10^7$	1:1 (neat)	not done			
$4.4x10^5$	1:100	38841	22	0	100
$5.0x10^4$	1:1000	38841	17	0	100
$7.0x10^3$	1:10000	38841	12	0	100
$6x10^2$	1:100000	38841	7	0	100
$3.0x10^1$	1:1000000	38838	2	3	40

ID_{50} calculations, based on the method of Reed and Muench {, 1938 #225}

Formula: $ID_{50} = [neat]cfu/200\mu l/10^{\log 10 endpoint}$

$\log_{10}endpoint = \log_{10}\chi$ - (proportional distance.$\log_{10}$ dilution factor between χ and y)

χ = dilution of neat media resulting in % morbidity directly below 50%

y = dilution of neat media resulting in % morbidity directly above 50%

proportional distance = (50%-X%)/(Y%-X%)

X = morbidity immediately below 50%
Y = morbidity immediately above 50%

Examples

Isolate #13 (assay E.1)

Proportional distance	= (50-20)/(80-20)
	=0.5
log10 of endpoint	$=\log_{10}100000-(0.5x\log_{10}10)$
	=4.5
ID_{50} (cfu)	$=1.2x10^8/10^{4.5}$
	$=3.8x10^3$ cfu

Isolate NAFC (assay A.1)

Proportional distance	= (50-0)/(100-0)
	=0.5
log10 of endpoint	$=\log_{10}10000-(0.5x\log_{10}10)$
	=3.5
ID_{50} (cfu)	$=8.0x10^7/10^{3.5}$
	$=2.5x10^4$ cfu

Isolate NAFC+BupsΦ1 (assay A.1)

Proportional distance	= (50-40)/(100-40)
	=0.167
log10 of endpoint	$=\log_{10}1000000-(0.167x\log_{10}10)$
	=5.833
ID_{50} (cfu)	$=5.6x10^7/10^{5.833}$
	$=8.2x10^1$ cfu

This calculation was programmed into Microsoft Excel for ease of use. Following are examples of the spreadsheet, along with the equations used.

Table A.4.4 Example of excel spreadsheet for use in calculating ID50 values. Yellow shaded squares are for data entry. Blue shaded squares contain formulas and produce numerical results identical to those calculated manually (previous page). Columns C and D represent the neat bacterial concentration $Wx10^{Z}$ cfu/200µl where W is placed in column C and Z is placed in column D. Data in columns E and G are dilutions in format 1:y, where y is the value entered, eg. 1:1000 enters 1000. Diln is short for dilution. '+50' represents the dilution at which above 50% of animals fit criteria. '-50' represents the dilution at which below 50% of animals fit criteria.

excel column	B	C	D	E	F	G	H	I	J	K	M
excel row				y	Y	x	X				
3	sample	[neat] cfu/200ul	by 10 expn.	diln +50 morbid	% +50 morbid	diln -50 morbid	% -50 morbid	diln difference +-50	proportional difference	log10 endpoint	ID_{50}
4	#13	1.2	8	10000	80	100000	20	'+G4/E4	'+(50-H4)/(F4-H4)	'+LOG(G4,10)-(J4*LOG(I4,10))	'+C4*POWER(10,D4)/POWER(10,K4)
5	NAFC	8	7	1000	100	10000	0	'+G5/E5	'+(50-H5)/(F5-H5)	'+LOG(G5,10)-(J4*LOG(I5,10))	'+C5*POWER(10,D5)/POWER(10,K5)
6	NAFC +BupsΦ1	5.6	7	1e+05	100	1e+06	40	'+G6/E6	'+(50-H6)/(F6-H6)	'+LOG(G6,10)-(J4*LOG(I6,10))	'+C6*POWER(10,D6)/POWER(10,K6)

APPENDIX 5

CALCULATIONS FOR STATISTICAL ANALYSIS OF BACTERIOPHAGE ASSAYS

Each assay and combination of assays is compared to the gold standard of all three assays, as described in Chapter 4. Data is collated from Chapter 4, Table 4.2. Data is presented in the following format;

	gold standard positive	gold standard negative	Total
assay positive	A	B	A+B
assay negative	C	D	C+D
total:	A+C	B+D	A+B+C+D

Equations used to calculate values are as follows (using letters as above to describe equation). All results are presented in the format of percentage;

Sensitivity (%)

$$\frac{A \times 100}{A+C}\%$$

Specificity (%)

$$\frac{B \times 100}{B+D}\%$$

Positive predictive value P.P.V. (%)

$$\frac{A \times 100\%}{A+B}$$

Negative predictive value N.P.V> (%)

$$\frac{D \times 100}{C+D}\%$$

True prevalence

$$\frac{\left\{\left(\frac{A+B}{A+B+C+D}\right) - \left\{1-\left(\frac{B}{B+D}\right)\right\}\right\}}{1-\left\{1-\left(\frac{B}{B+D}\right)\right\} - \left\{1-\left(\frac{A}{A+C}\right)\right\}} \times 100\%$$

Accuracy

$$\frac{(A + D) \times 100\%}{A+B+C+D}$$

Data is as follows;

Comparison of plaque (P) assay to gold standard to test for bacteriophage

	all assays positive	**not all assays positive**	**total**
P +	11	12	23
P -	0	8	8
total:	11	20	31

sensitivity	100.00	%
specificity	40.00	%
P.P.V.	47.83	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	61.29	%

Comparison of digest (D) assay to gold standard to test for bacteriophage

	all assays positive	not all assays positive	total
D +	11	5	16
D -	0	15	15
total:	11	20	31

sensitivity	100.00	%
specificity	75.00	%
P.P.V.	68.75	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	83.87	%

Comparison of mitomycin C (M) assay to gold standard to test for bacteriophage

	all assays positive	not all assays positive	total
M +	11	2	13
M -	0	18	18
total:	11	20	31

sensitivity	100.00	%
specificity	90.00	%
P.P.V.	84.62	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	93.55	%

Comparison of plaque (P) and mitomycin C (M) assay to gold standard to test for bacteriophage

	all assays positive	not all assays positive	total
M & P +	11	1	12
M & P -	0	19	19
total:	11	20	31

sensitivity	100.00	%
specificity	95.00	%
P.P.V.	91.67	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	96.77	%

Comparison of mitomycin C (M) and digest (D) assay to gold standard to test for bacteriophage

	all assays positive	not all assays positive	total
M & D +	11	1	12
M & D -	0	19	19
total:	11	20	31

sensitivity	100.00	%
specificity	95.00	%
P.P.V.	91.67	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	96.77	%

Comparison of plaque (P) and digest (D) assay to gold standard to test for bacteriophage

	all assays positive	not all assays positive	total
P & D +	11	3	14
P & D -	0	17	17
total:	11	20	31

sensitivity	100.00	%
specificity	85.00	%
P.P.V.	78.57	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	90.32	%

Comparison of plaque (P), digest (D) and mitomycin C (M) assay to gold standard to test for bacteriophage

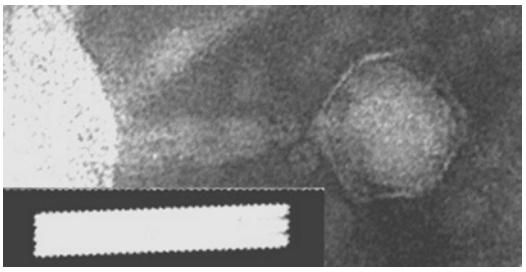
	all assays positive	not all assays positive	total
M & P & D +	11	0	11
M & P & D -	0	20	20
total:	11	20	31

sensitivity	100.00	%
specificity	100.00	%
P.P.V.	100.00	%
N.P.V.	100.00	%
true prevalence	35.48	%
accuracy	100.00	%

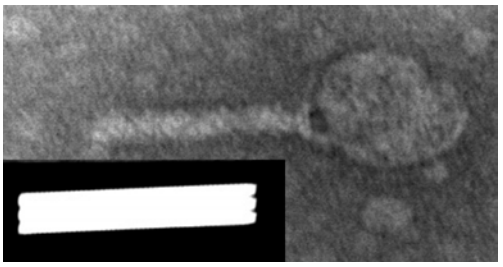
APPENDIX 6

TRANSMISSION ELECTRON MICROSCOPE IMAGES OF BACTERIOPHAGE EXTRACTED FROM *B. pseudomallei*

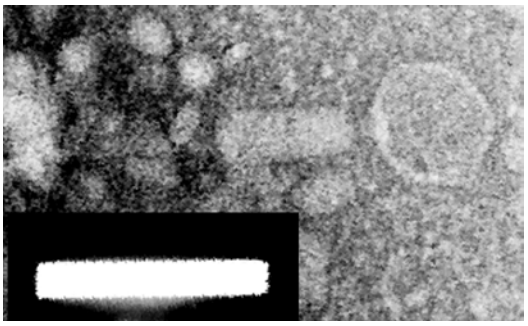
The name of the isolate is noted at the bottom each image. The size bar in each picture represents 100nm. Images are representative of those observed for each isolate. All images were taken at the Centre for Microscopy and Microanalysis at the University of Queensland and the image trimmed and resized using Corel Photo Paint 11®.



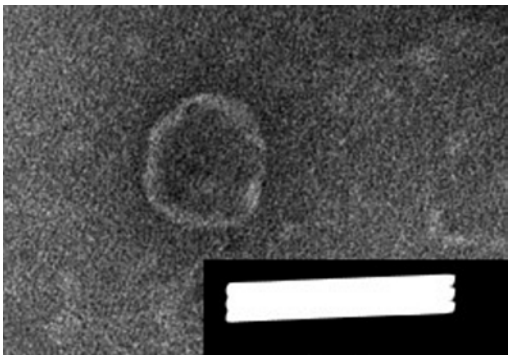
Isolate NCTC 13178



Isolate NAFC

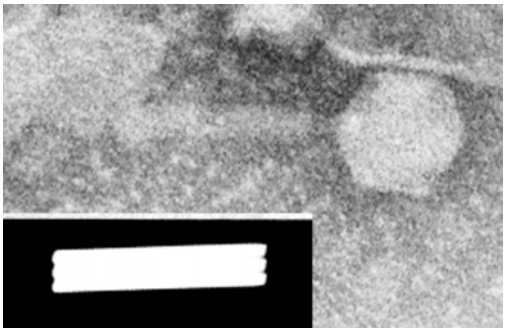


Isolate #2



Isolate #7

Isolate C4

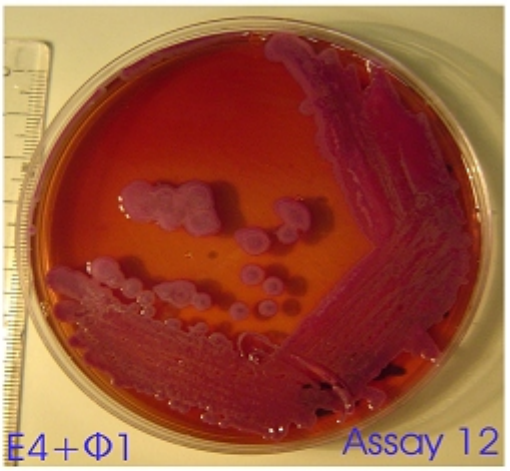


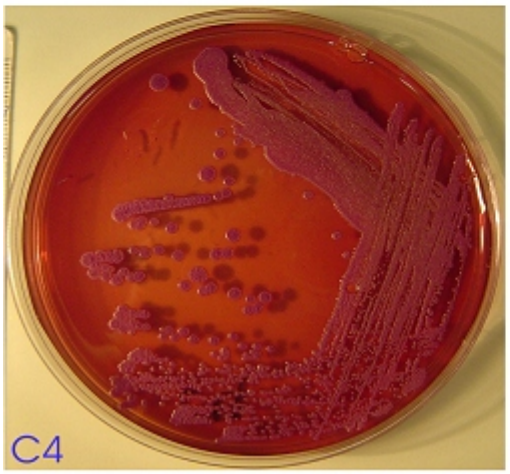
Isolate *B. thailandensis*

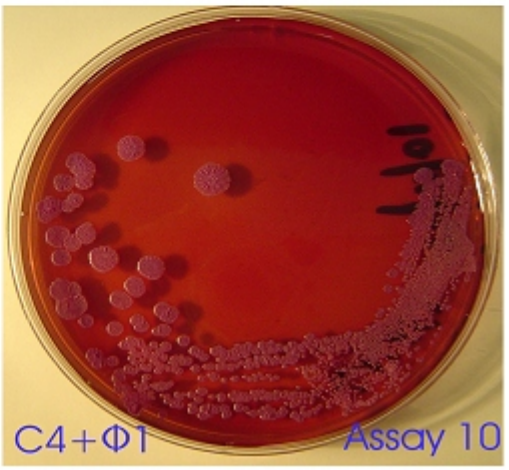
APPENDIX 7

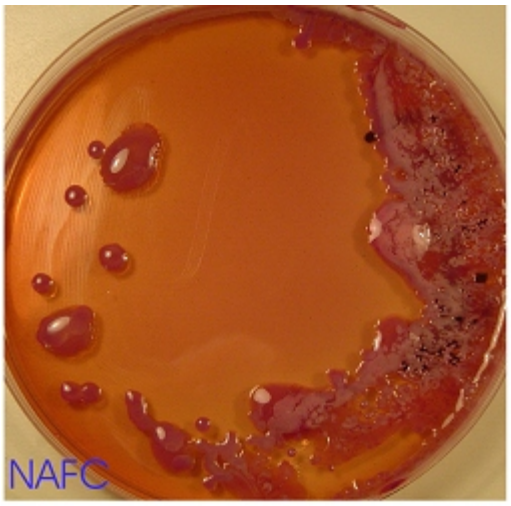
IMAGES OF COLONIAL MORPHOLOGY OF ISOLATES BEFORE AND AFTER EXPOSURE TO BACTERIOPHAGE

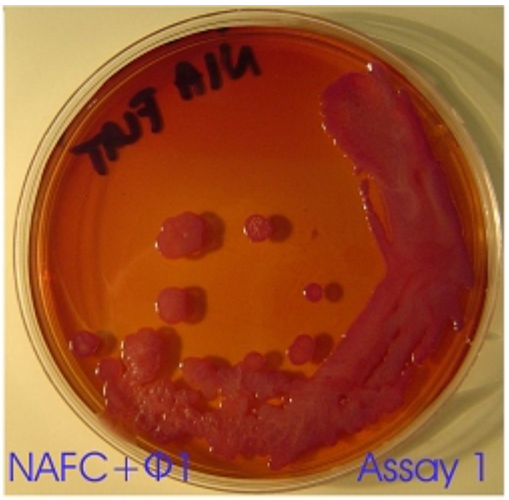
The name of the isolate is noted in the lower left hand corner of each picture. If exposed to bacteriophage (Section 5.2.4), this is also noted in the lower left hand corner and the assay in which exposure took place is noted in the lower right hand corner. All colonial morphology are shown on Ashdown agar (Appendix 1). All images were taken with a digital camera and the image trimmed and resized using Corel Photo Paint 11®.

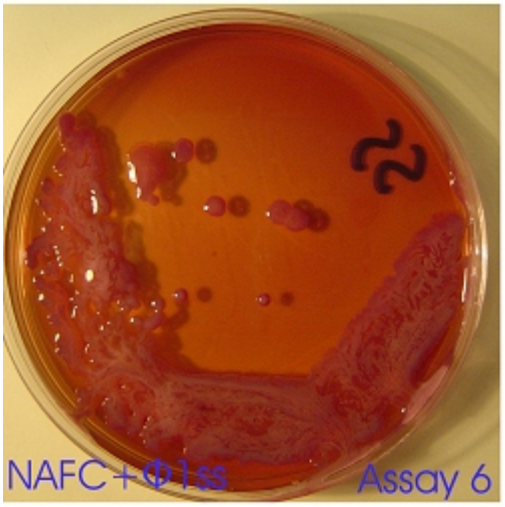


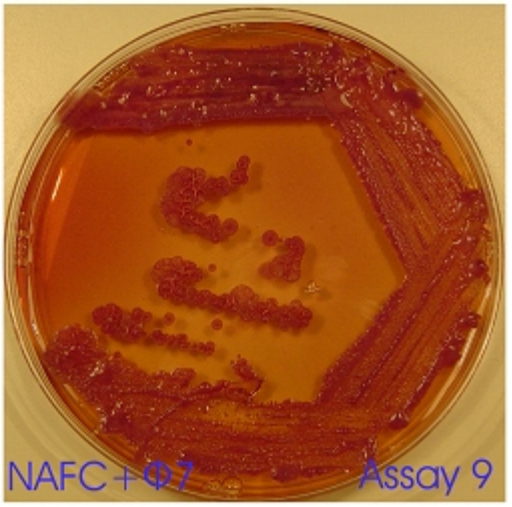


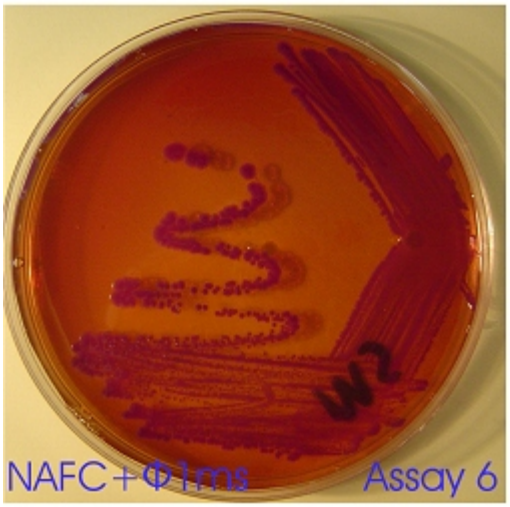


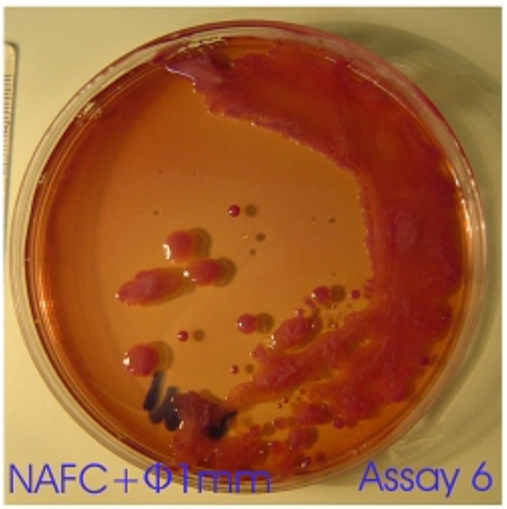

NAFC

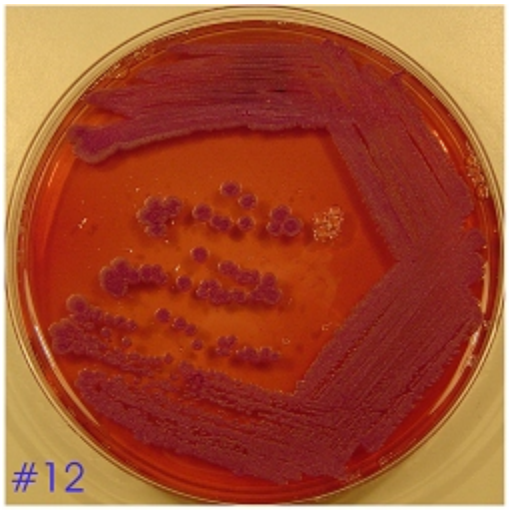

NAFC+Φ1
Assay 1

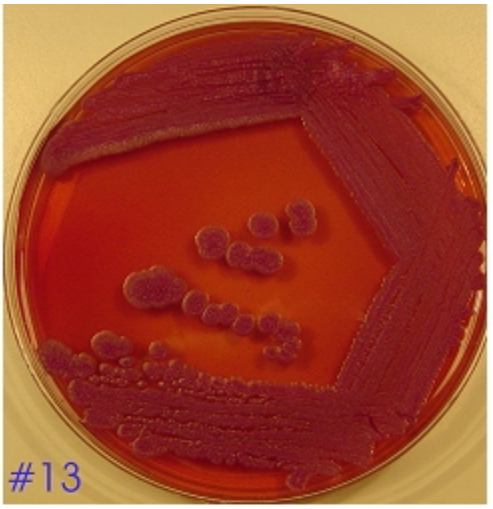

NAFC+Φ1ss
Assay 6

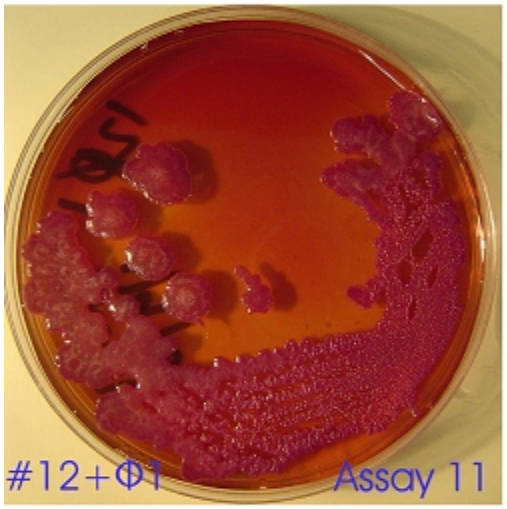

NAFC+Φ7
Assay 9


NAFC+Φ1ms
Assay 6


NAFC+Φ1mm
Assay 6


#12

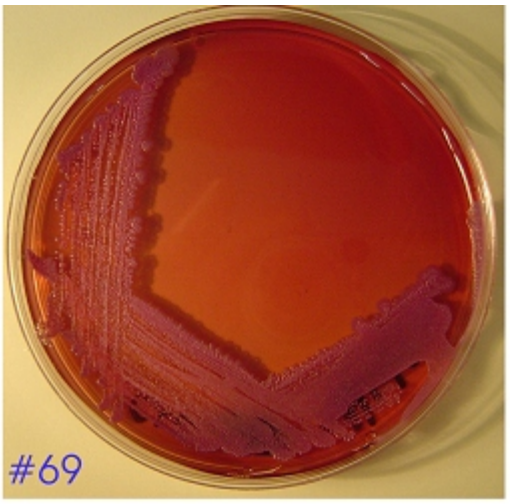

#13

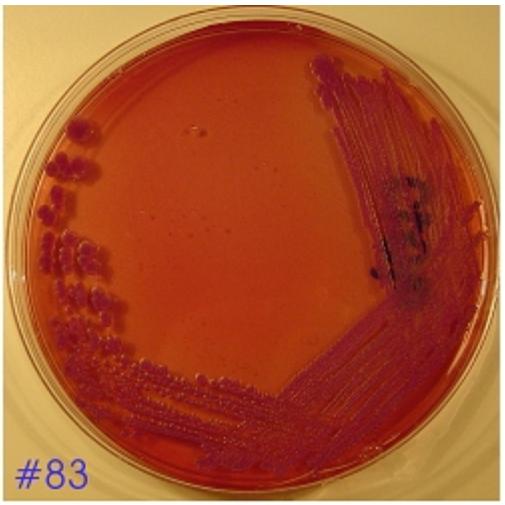

#12+Φ1
Assay 11

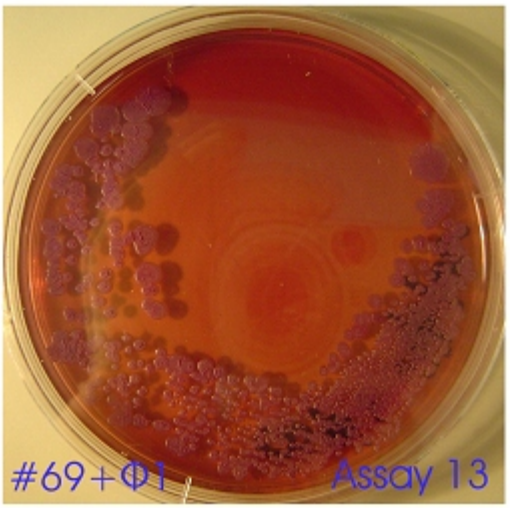

#13+Φ1
Assay 15


#12+Φ7
Assay 11

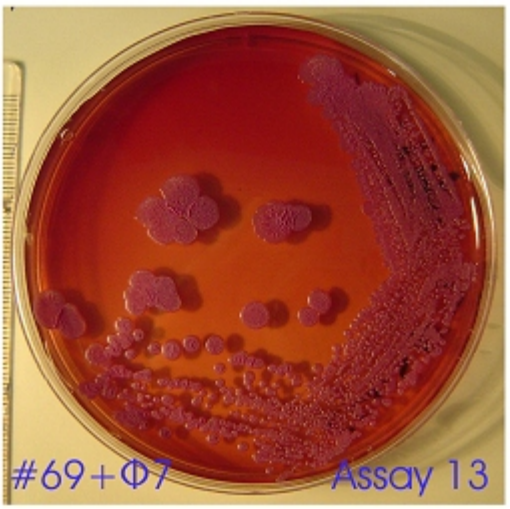

#13+Φ7
Assay 15

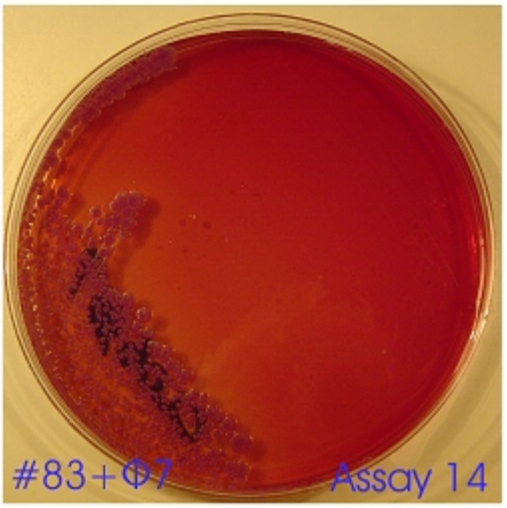

#69


#83


#69+Φ1
Assay 13


#83+Φ1
Assay 14


#69+Φ7
Assay 13


#83+Φ7
Assay 14

APPENDIX 8

SEQUENCE OBTAINED FOR BupsΦ1; AS ANALYSED IN CHAPTER SIX

Contigs are placed in the same order that they are placed in Table 6.3

Figure A.8.1 Contig C 5846bp

ATGGGCGGCG GCAGCGTTAT CGAATACATC GGAGAGCGAC GCGCAGAATA CGCGCTGGCG 60
GCGGTCGTGC TTCGCGAAGA ACTCCCCATG ACCTGCATTC CATCGCTTCG GATTGTTCCA 120
GGTGGACGCG GCTGTGCGAT GGCGCGGATT GCCGGCTCCC CAAACCACGC GAAGTGCCGT 180
GCTGGCGCGC TCGCGCTCGG CATAGCAGTG GTCGCAGCCG GGAGATATCT TCGTGCAGCC 240
GATCCACGCA TTGAACGTAT GGTCCGTCCA TTCGATTTTG CTGTTCTCGC TCATGGTCTG 300
TTGTCCGCGC CGCAGTTAGG GCAGTCGAGG GAATAACGGT CAGCGTCGCG CAGGAAGCGG 360
CCACATCCGA CGCAGTTCAC GCGGTCGGCG TGCTGCTTCC GCGGCTTGGT CAGCGTGATT 420
CCGGTGCCAG CGAGCGCCTC ATCGCGCTTG ATGTGCTGGG CGTCGACGGC CGGCCGCCAC 480
TTCGAGTCGA TGTAATCCTT CGGCCACGGG ATATCGGTCG CGCGCGCGTC GTGGCGTGCC 540
TGCGCTTCTT CCTGCGTATA GACGTGCGCG TTTCGCAGGT CGGTCGTATA GCCGCTTCCG 600
TTCCGTGCCC ACCACAGCAC GTCGTTGCCT ACGAACGAAC GGCTGTCCTG AAGATAGAAA 660
AGAGCGCTCA TATCTGGATG TCTACGTATT TGGTGATGAA CCGCGCGTCC TCGCTGTGGT 720
GCCGGCGATA GACGATACGA ATGCCGCCAC CGAGGTCCAG CGGTCGGCCG CGGCCGTTGT 780
CGTCGTTCAG CTCGGCCTGA CGGCCGCAGC AGGCGGAGCG CAATGTGTCG GCGAGGCCGA 840
GGGGATCGGC CAGCAGTCGA AGCAGGAAGT CGTGGTTGAT CTCAATGATG GTTCGCGTGC 900
TCATGCTAGG ATTGCCTCAA CATTCGAAAG GAGAGCGCCG TGGCGTACGA ACCGACACTA 960
GGCACGGCCC CGGTTGTGAT CGGGGATGAC GTTCAGTTTT CGCTGACAGT GGACGGCCAC 1020
GTCCAACGCT TCAGCGTGTC GCACGAAGCG CTCGAGGATC ACTTCAGCGA TCCGAACGGC 1080
AGCGCGTTCG ACCGATTGGA AGCGTTCGAG CGTGGACGCG ACCGGATCTA CGCGGCGGCG 1140
GCCGCAAAAC TTGGTATCCG ATCCGGCAAC GTGATCGGCG TTGGCACACA CGATTTTTGA 1200
CCTGTACTAC CGATGAAGTC TGTTTTCCTA TTGCGCGGCT ACAAAGTGAA TTGCACGCCA 1260
CGCCCGACCG ACGACGGCAA GTTCGCCGCG CAGGTCGAAG TGACGAAGGT GGGATTCAGC 1320
CGCGAGGCGG CGTTTCGGAA ACTCGGCGAG TTCGACACAG AGGCCGAAGC AGTCTCCTAT 1380
GCGAAGAATT TTTCGGAAGA GTGGCTGACG CGGTACGCTT AAGTGCCCGG ATCACTCGAG 1440
CGAAGATCAA TCAATTTCCC GGGGGTTGGA TATGCGCAGA GGCATTGAGC ATTTCATCGA 1500
CGAGTTCGAT GGCTATCAAA TTGTCGTCGA TTGTGAGCAG CCGGCGCCTA ATAGCTCTTG 1560
GGTTCTGAGC GTTCAGATCT ACAGGGATGG GGAGCCGGTG CTTCCGCGCT GCGATGATTC 1620
AGACAGGGCA TATGCCACTG CGGAGGAGGC AAAGCGAGTA GGTGTCGGGC TCGGGCGGAA 1680
CATCATCCGT AATCTCAGCC GACAGTAGTT AACGCCTTCC ATCGCCGGCA GCATGCCGTC 1740
GGTGACCTTC ATGCTCGCAC CCATCCTTTC GACGTCGAGC GGATCTGTCC GGTCTTGCGC 1800
ACCGCCTGAA GGCACCGATC AACAATGCGC CACCCCACGA CTTCACCGAA TGTGGTCGGC 1860
CGGCATTCTT CGCGCGCGAG TCGCTCGCTT TCTTCACGCA CTGCGCCGGT GTTGACGGCT 1920
GCGAATTTCT TCGGAGTGTC GTCGATCGAC GCGACGATCA GGTCGTCGAG TTTCTTGTAC 1980
TTGCTCATGA TGGGTCTCCT GAATACCGAA CGAACATGCG GCGCGGGCGC TCGCATGTCG 2040
GGCAGAGGTT GACCTGACAC CATTCTTCGA CTGGCTCTGG AGCGTGAATG TGCACGAAGT 2100
TCTCAGCCAT TGCCTTCCTC ACTCGATGCG GCGGTCGGAG TGGTCTTCAG AGCGCGAATC 2160
TTCTCGGCCT CGTCACGATA GATGCGCGAT ATCGTATGCG CGATCATGTC CCCGCACGCT 2220
TCGATCTCGT CCGCGCGTTG CTCCATGAGC ACGGCTGCTT CCTCAAGGCC CGCGTCCCTG 2280
TCGGTAGTGG TGCGGCGGGC TGCGAACGTA CGCTGCCAAG TTCCAACGCA GCTGCATAAC 2340
TCGCACCATC CATCATCGTG CCGCGCTCGC GGATCCATTG CGATACGATC GGCTCCATTT 2400
CCTTCAAGCA TGTCATCGGT GATCTTCAAC TCTCATTCCC TTGATAGGCA ATTCCGATTT 2460
AGCCTGCCAC CGCAGTAGCA GCAGTAGCGC ATGCCGTTTT CCTGCGGCAT ATCGTCGAGC 2520
GTGACATCCT TACCGCAGGA CGTTTCCCAG ATATCCGACA TGCAGTCTGC GGGCGTCCAC 2580
GAGCACGTGC GCTCTGACGA CTGTTTTCTC CCTTTACTCA TCGCGCCCTC GTCGCCATAG 2640
CAAGCTCGCG GCGGCATGCG AATTCGCGAT CGTATCCGGC AAGCGTCGAT TGAAGGTTCG 2700
GGCCGGTGCG TCGCTCACCT ACCGGTGTAT CGAGCGTCGC TCGCGGCGGT GGTTCGGCCG 2760
GTGCGAGGCA ATAGCCAAGA TCATTCGTAC ATCGCAAGTC GATACGAACG CGGCCCGAAT 2820
AAAAGAGCAT GTCGACTCGA CATTGAAGAA CGGCAGGCTG AACGCCAAGC GCGAATGCGA 2880
GCTGATAGAT GGTCTTCGTG CCGCTCGTGA GCTGCGCGAG GATGTCGTCG TTCTTGATGC 2940

GGGGAGTGGT CACGATTCGC CTCCTTGGGT GCGGGCGGCG TCTATCTTGA GATACTCGGC 3000
ATACTCTGGC CAATATTTCG CGATGCTTTC CGTGTGTTCG GCTGCGGTTC CATCCCATTC 3060
GATAACGACA GCAGTAACAG TCGGCCAGTG GTCTGGATCG TGCCCATCAG CAACCGACTG 3120
TGCTCGTTGC TTCGTGAGAT ATTCTCCCCA ATACGTGTTG AATGCGGCGG CAACAGCATC 3180
GGCCTCATCC TTCGACGGCG CGGCCACAAT ATCGTCCGGG CCTTGAATAT TGAGCATCCA 3240
GAGTTTGGTC ATTGCTTTTC TCCCGCTCGG GCGGCGTTGA TTAGCCGATA CACTTCCTGC 3300
AATTCCGAGG CGCGCTTGAT CACACCATCA CGGTTGATGA TCGCGAGCAA CTCGACGACG 3360
CGCGCCGGCA CACGCACGTA GTCCGTCACC TCTCCGCACG GCTGCTGCGT CGTGAGTGCA 3420
TCCTGTCGGC ATCCGTCATA CTGGCAAGCC CGAGGCGTCG AACAGGCCGC GCACGGCATG 3480
TTCCAGTCGC TCAGCTGCTG CTCTATCCGC TGTTTAAGGA GCGCGATTAT TTCCTCCGAC 3540
TGAGCCGCGA GATAACGGAT TTGGTCTTCA CGCTCGCGCT CGTAGAAGGC GAGCTTGGAC 3600
TTGAGTTCGT CGATTTCGGT GCTCGCCTGC TGCGTGGCGG CTAGAAGCGC GCGCATCGCT 3660
TCGAGGTGGG TCCGCGACAC CGCCATCGCT CTCCGGTCAT GCGCTGTGAT CTCGCGATCG 3720
TAGGTCGCTG CCAAGGCGCG GAGCGCGGTT TGCTGGGATT CACCCTTCTC ATTCAATTCG 3780
AGAAGCTCGC GGCAGCGTTG CAGCAAGTCG TCGGTGGGCT TATCGGTCAT GGTCGGCCTC 3840
TTGCAATGCC TTGATGAATG CGGCCGCCTC TTTCAGGGAA TCGAACAAGT AGGCCTCGTT 3900
CGTATCGCCG TTGTAGACAT GCGCCCACGT TTTGATCTCG GGCTCTCGAA CTTGCACGCT 3960
GTACTTCGGC TTGAGCGTGT TCATGTCAAG CCACGGATGA GTGCGCATTC TCCATTCCGG 4020
GCGCACGCTG CTGCGCGCGA TTCCAAGTTG CGTCTTAGCC ATGGCTGGCT CCCTCGGCAG 4080
TTGACTTGCC GCAAAACGGG CAGTAGCTGG AAATGACCGG AATCAGCTTT CCGCGCGTGA 4140
ATCCTTTCGC CTGCGCGACG ATCTTGAACT CGGTCTTGTG GATCGCGCGT ATCGAATTGC 4200
CAGACATGGA GAATCCGGCC GATTGGCAGT CGGCCGAAGC GTCCACGCCG AGTTCTTCGC 4260
TGTAGCGTTT GGCCAGTTTG ATCTCGATCT CGCTTATGCA ATTGCAGTTC ATGATTTGTC 4320
GGCTCCATTG AGAAGGGCGC GGAGCACACA GGCATGGGCA TTTAACATGC GCGCCTCCAT 4380
TGTCTTGGCT GCAAACTCGA TTGCCTCGCG CTGGCTGTTC GTCAGGCTCG CCACCCTCGC 4440
GGCGGGCGGG GCGGTGTAAA GAGGTTGCTT TTCGTAATCG CGGTCGCAGA CATGGACGGG 4500
GTCTCCATCG CAATAGCACC AGTTATCGCC GCTCGTTCGA TAGCGCCACG CTGCCGCCTC 4560
TCCCGCATCG GCGGGGGCGC TGACGGCGTA GCACGGAATC GAATACGGCC GCACCGATGA 4620
TGCGGTAGCG CCACCATCTG CCAATGCGCG TTGCTTTTGC GCGGCGGTGA TTGCCCGGTC 4680
GTCATCAGTT ACCCACGCGA TCGGCTCGCG CGCCTCTGCC GGTGCGTCGG CTTTGTCCAA 4740
AAACGCATTA GCGACTCGGA CAAGTCGATT GAGATGATCC TCAATCGATG CGCTGAAGCC 4800
TTCGATGGAT TCATCGGCGG CAATGCCAAA CCGCGTCATC AGCTTGCGCA CGCGCGTCAA 4860
GCAATCGACG CCCACATCGA TTACTGTCGG TGCGTCGGCC TGCGCTGCAT CTGCCTCGCC 4920
GTTCAATACG GCATTCGCGC GATCAACCAT GCGCACTGCC CACGAATGGC CGCCCGGATC 4980
GGACTTCTTC ACCCAATTCG ACATGACGCG CATGGCATCT CGTGCCTCGA ACAGCCCTTT 5040
CCGCAGCGCT TCGATTCGAA CAGGTGCGTC TGCCTGCGCG GGTTGCGGGG TACTTTCGAA 5100
AGCATCCAGA CGCTCCCCCA TTTCAGAAAG CAACTGTTCA ACGTCCGCTA CGTTTCTTGC 5160
GTAGACTTCG TGACCTTCAC CATCCATCCC CGGAACAACG TTCAAGAGAA TCTTTCGCAC 5220
GTCGCAATCA TCCGGCACCG CCTCCGCAGC GGGCGATGCT GCCGCCAGTG CACGTTCCGC 5280
CTGCGCCTTC ATCCGATGTG CCACGTCAAT TTCGCGCTTC CAGTACCCGG CCGTATCGCT 5340
GTCTTGGGCT TCTTCGGCAG CGAGGCGTTG AGCGGCGAAG CCCGCCACGA TGTTGTCATG 5400
GTGGCTTACC ACGTCCATTA GGCATTCGCG CGCGTCGTCG AGTTGCTCGC GATCATCAAC 5460
CGTGATGGTG TGATCAAGCG CGCCTCGGAA TTGCAGGAAG TGTATCGGCT AATCAACGCC 5520
GCCCGAGCGG GAGAAAAGCA ATGACCAAAC TCTGGATGCT CAATATTCAA GGCCCGGACG 5580
ATATTGTGGC CGCGCCGTCG AAGGATGAGG CCGATGCTGT TGCCGCCGCA TTCAACACGT 5640
ATTGGGGAGA ATATCTCACG AAGCAACGAG CACAGTCGGT TGCTGATGGG CACGATCCAG 5700
ACCACTGGCC GACTGTTACT GCTGTCGTTA TCGAATGGGA TGGAACCGCA GCCGAACACA 5760
CGGAAAGCAT CGCGAAATAT TGGCCAGAGT ATGCCGAGTA TCTCAAGATA GACGCCGCCC 5820
GCACCCAAGG AGGCGAATCG TGACCA 5846

Figure A.8.2 Contig AB 13865bp

CTTGTGTGAT CCACGTAGTG TCTGACCAGC GCGGGCAGCA CACGGACGTT GTTGTGATGG 60
GTTCGAGGAG GGTGCATCGC TTCCCCTTTA ATAAAGGATT ACTAAGAAAA TTGCAGTCTT 120
GAACTGCGAT GCTGCATGCG CAGCTCTCAC TGGTGACGCC TCGAGAGTAC GAACCGCCAC 180
CGGTCAGAGC TGTGAGAAGT GCCGGTTACG TTATCCGGCG CCGGTTAGCA TCCCGGCCGT 240
CTCTTTGAGC GACCTTCCGT CGGCTTGCCA AACCTGCTTT CTGAGGTGAC GGGCCCACTT 300
CGTTTATCCC TCGCAGGGCG GGCGAGTCGC GCTCACCGAT GGAGTTGGAT GTCCATCTAC 360

CCGGGCGCAT TTTCGAGGCA GGCTCATGGA GCCGGGCCGG TACTGATCTC CGGCTTCGGG 420
ACTCCAGCAT TGAACTGGAT GCGCATCAGC CTGCGCATTC CGGCTCCATG AGACTGCTAC 480
CACTCGCGCG CCCGGCTACT CCCGGCCGTG CCGGCTCCGG GCCGCGCGAG GTTTGTGCCG 540
ATTACAAAGC CATCGGTCAC GTGGTGCTGG CTGTCTTGTG TCAGGTCCGT TCAAGCCTAC 600
AAGCGGTAGC CAAACATCGG CCTAGCGCGC TGCGCCTGTC CTGACTCACG ACGCTGATCG 660
CGCCGGCCGG TTGCTCCGCC GAAGCGGTCC GGCATACCTT CGATTGTTAG AGAGCGATCC 720
GCCTAGGGCG GTGGCGCGGC ATCGGTGTCG CGTTGATTGG AATTATAACC AAGGTTATCG 780
TCGTGTCAAC AACCAAAGTT ATCATATGGG CGGGAAATTT GTAACAGCGC GGCTTCGGGC 840
GGTTTGTTAA GCGTCTGCTG ACGGCAGTTG AAAGTCTTCC GGAACTGGGG CTAAACTACT 900
GTACATACAT ACAGTATTGT GACAAACCGA AGACGAGGGC GGCCAGTGAA AGAAGAATGG 960
AAGACGCGCA TGCGCTGCAG GCCTGGGGAC TTGGCTAGGG TGGTGGCGAG CAGCAATCCT 1020
GCCCTGATCG GCACGATTGT GACGATTCAG CGTATGCGGT CTGACTATCG ATGGGACGTG 1080
CTTCTTGAAA CGCCAGCGTT CGGAATCACC GAACGAGCAA AGCGACCCGT CGTGACGCGT 1140
GAGTTCTCAT TTTGGGATGC GTCCCTGGAG CCGCTACCGC AGAGCGCGAT CTTTGTCAGT 1200
CGCCGCGCGG CCTATCTTCG TCCTCGGGAG AACGAGGGGC GGGAAGCGTT TGGGCTAGCA 1260
GGCCACTGAT ATACGCCTCA ACCTTGGCGC GGCCGAGGTC GTCGAGCTGA TCGTATCCGG 1320
CTGGGATCTG TGGGCCCACT TCGTCGCCGA ATTCGAGCCA CTCCAACGTC GTCAGCAGCG 1380
CTCGCGCGAG TGCGAACCCA TTTGTAGTCG TGGTGCTGCC GCCTCTTTCA ACCTTGGCGA 1440
TTGCCGGCTG CGACACGCCG ACAAGTTCTG CGACGTGCTT CTGGGATAGG CCCAGTGCCT 1500
CACGACGTTC CTTTGCCCGG CGCCCGAATT CAGGATTTTC CATAGCGGGC ATCGTATAAC 1560
CTTGGTTGTA GGTCGGCAAA TAACCAAAGT TCTTGCGAAT AAATAACTTT GGTTATAGTA 1620
TTGCCGCATG AAACCGACCA ATTGCTCCGT GCCGGCTCTG CAGGCAGCGA TCCAAAAAGC 1680
TGGGTCGCAG TCCGCACTTG CCCGCCTGAT CGGCAAAAAG CAGCCGCACA TCCACAAGTG 1740
GCTGCATTCC CCCAATGCAA TGAGGCCCGA GAACTGCGTC CTTGTCGGGA ACGCAGTCGG 1800
CATCCCGTAC AGGGACTTTC GACCCACCGA TTGGCATCTG ATTTGGCCTG ATCCCGTGGC 1860
AGGTGCCTCG CCGGAACGCT TGGACGAGTG CACGTTGGAG GGAGGGTGGC GCGACGCGCT 1920
GCGTGCAGGC GTCGACTGAT CGCAATGGTT CGCATTGTTT GAAGGCTGGC CGGGTGTTTT 1980
CCTGGCCTTT ATTTCGCCCC GGTGCCAACT GGGTAAGCAA GTGGGTAATC AACTGGGTAA 2040
CGATTGATTT TTCGTATGAA CCAGACCGAA TTCAGGATGT TCGCGCCGTG GGTGCAGGCG 2100
GCGACACTGC CGGACGGCGA GATCGAGGCG ACGAGCTTCG AGGACTGCCT CGCGCACGCG 2160
CTCGAGCTCG GCCTGCGGCG CTTCGATCGC AAGACGCTCG CGCTCTACTG CGACATTCAC 2220
TATCCGCACT TCGGGGACCT GATCGCCGGG CGCCGGCCGT TCCCGGCCAC GAAGCTCGAC 2280
CGCTTCTGCA TGTTCACCGG CTGCGATTAC CCGCGGCAGT GGCTTGCGAT ACAGGAGCGC 2340
AAGGCGATCG AGGAATATCG GCGGCTCAGC CAGCAGGCGA TCGGCGAGTT CGTACAGCAG 2400
GCGTTCGGCC AGCGGCAGGC GGTGGCATGA CAGTGACGCT CAGCAACCGC GACGTCGGTA 2460
AGTCCTTCAC CCGCAAGCTC GGGCGCCCCA TGACTTACCT CGGCCTCGTC GAGGAAAAGC 2520
ACCTTTTCAT CTTCCGCGAT CCCCCGCAGG ACTACCTCGC ATTCCGTCCG GATCAGCTCT 2580
GGATGCTGGA ACGCATGCGT CCCGATGCGG CGCCTATCGA CAGCAACAAG AAGGAGGGCG 2640
GGTTGTGCTG ACGCATCTGT TCGAGCGCGC GGCGTTCCGC GCAGGGTGGC GCGCCGCGCG 2700
CGCCGGCGTT CCTTTCCACG AAAACCCGCT GCGTGGCGCG CTCGCGTGCT TCGCTCAGCA 2760
GTGGGGGCGC GGTTGGGCCG CAGCGAACGA CTTCCCGCGG CCGTACACCT TCCACGACTG 2820
GGAGCGGTGG ATCTGCACGA CGTCCGCTGA ACAGTTTCGG GAGGCAGCAT GAAACGCCCC 2880
TCGAAAGCCG CGCTCGCGCG CGAGAACGTC AGCCGACTGG CGTCGATCGG CATCGCTGCG 2940
CTCGAGTACG ACCATGCACG TGTCGACGCG AATGCTGGGC GCAGGTCGCT GATTGCTGCC 3000
CGCAAGCTCT GGAGCGAGGA CCACGATCGC AGCGACGACA CGCCCGCCGA GTGCGAAGAG 3060
CAATACGCGC TGTCGGCGAA GAAGCGCACG CGGGCTCGCG CAAGACTGCT CCGGCAGATC 3120
AAGCGGTACC GCGAATGGCT CGCAGAGGTG TCGGCATGAC GTGGGCGCAC GACGACCTCG 3180
CCAAGGATCT CGCCGCGCAT CTGCGTGGCG CGTCCAATCG CCTCATATGG ACCGACATGC 3240
AGCTCGGGCC GTCCGGTTCG CCGCGGCCCG ACGTCTATTC GGTGCCGTGC TCATTCTCGC 3300
GTTTCCAGCC GATAGCGTAC GAGTGCAAGG TCAGCGTCGC CGATTTCCGG CGCGACGTGA 3360
CGACCGGGAA GTGGACGTCG TACCTGCGCT TCGCCGCTGG CGTGATTTTC GCCGCGCCGG 3420
CCGGCATCCT GAAGAAGGAA GACATCCCAG CTGGCTGTGG ATTGACCGTG CGCGGCCCGG 3480
ATGGCTGGCG CTCGCTGAAG GGCCCGACGC TGAAGAACAT GGAGAACCTG CCGCGCGACG 3540
CATGGATCAA GCTGATCATC GACGGCATGG CGCGGCTCGC GGATCAAAAC CACGAGCAGT 3600
TGCGCGCGGG TCTGTGCAAC GAATGGACGC TCGAAAAGAA GCTTCGCGCG CGCCTGGGCG 3660
ACGTCGTCGC GGACGCGGTC CGGGACCAAC TGCACGCGGA GCGCCGACTG AAGACGGCCA 3720
CCGAGCGGCT TGAGAACCTG GCCGAAGAGG CGGAAAAAGA GCGGCGGCTG ATTCTCGATC 3780
GAGCGAAGGA GCATGCAGAG CGCGATGCTG CGCAGATCGA TAGCGCGCGC ATCGAGCTCG 3840
CGCGTGCGTT TGGGCTGCAG CCGAGCGCCG GCGCGTGGGA GATCGCCAGC GCCTGCAAGG 3900

AGGCCGCCCG CCGGCTCAGC ACTGATGCCG AAGTGAAGCG GCTCCGTCAG TTGATCGAGC 3960
GCATCCAGAT CGCCATCGGA ACGGCCGCCG AGCCGCTGCC GCATATCGCG CAGGAGGCGG 4020
CGTGAACTGG CTCGATTGTT CCCACCGCGG CGACTGCCGC GACCTCATGC CTGCGATGAT 4080
CGCGGACGGC GTGCGCGTGC AGACGATCGT CACGTCGCCG CCGTACTGGG GCCTTCGCTC 4140
GTATCTGCCT GACGGACATC CCGACAAGGG CAGCGAGATC GGCAGCGAGT CGACACTGCG 4200
CGAGTTCATC GACACGCTCG TCGGCGTGTT CGAGCTCTGC CGCCAACTGC TCGTGGACGA 4260
CGGGACGCTC TGGCTGAACA TGGGCGATGC CTATGCCTCA TCGGGCGGAC AGACGCCGAT 4320
GCGCGGAGAG ACGTTTGCCG GGCGCGCTCG CGCTAAGGAG AACATCTGCC TGAGCAACAG 4380
GAAAGCGGGC ATCGACGGTC TGAAGGTCAA GGATCTGATG GGCCAGCCGT GGCGTCTTGC 4440
GTTTGCATTG CAGGATGCCG GCTGGTATCT CCGACAGGAC ATCATCTGGC ACAAGCCGAA 4500
CCCGATGCCC GAGAGCGTGC GCGACCGCTG CACTAAGGCA CACGAATATC TGTTTCTGCT 4560
TTCGAAGAGT GAGCGCTACT ACTACGACTT CCACGCGATG CAGGAGCCTG TGAGCGGTGG 4620
TGCTCATGCA CGTTCGCCCG GCAATCGGTC ACACAAGGCC ACAAGTGCAT TTGCGGCGGG 4680
CGACGAGCAT CACCGCACGA AAAGCGGACT CGTCGCGTAC GCCGAGCGGC AGCGCGCCGC 4740
GGGCGTCAAT CGGAAAGCTG TAGCGGTCGC CGGTTGGCAG ACTGGACCGG GCGCACATTC 4800
GACTGTCGAG CACAACCGCG GCGCTCGTGC AAAGCGGCAG AAGCAAAACG AATCGTTCTC 4860
GGCAGCCGTC ACCGACGTTG TCACGAGTCG AAATCGCCGG AGCGTCAGGA CGATCCCGAC 4920
GCAGTCGTTC GACGGCGCCC ACTTTGCAAC TTTCCCCGAG GCGCTCGTCG AACCTTGCGT 4980
GCTCGCCGGC AGTCGTCCGG GCGACGTCGT GTTCGATCCG TTCTTCGGCA GCGGCACGAC 5040
GGGCCAGGTC GCGCAACGGC TCGGCCGCCG GTTCCTCGGC TGCGAGCTGA ATCCCGACTA 5100
CGAGCCGCTG CAGCGCGATC GCCTGCGGCA GCCGGGCCTG ATGCTGGGGG CCTCGTGAGC 5160
GTGCGGCCAA CCCTCCATGT CGTATCCCTG TCCGGCGGCA AGGACAGCAC CGCGACGCTG 5220
CTCGTCGCGC TCGAGCTGCA CGGACACGAG AACGTCCGTG TCGCAATGGC GGATACAGGC 5280
AACGAGCACC GTCTTACCTA CGAGTATGTC GACTATCTCG AAGACGTCCT GTCGATCCCG 5340
GTGGCGCGTC TCAAGCGTGA CTTCACTCCC GAGTGGTGGC ATCGGCGCGA CTACGTTCGC 5400
GACAAGTGGC CGGAGAAGGG TGTTCCTGAG AACGTCGTCC TGCGCGCCTT AGCCGTGTTC 5460
GAGCGTGGGC CGACCGGCAT CCCCTTCCTC GACCTTTGCA TCATCAAGGG GCGCTTCCCG 5520
AGCCGCATGG CTCAGTTCTG CACGTACTTC CTCAAGACTG AGCCTCTGAA TGAGTACGCG 5580
CTGAACCTGA TCGATGAGGC GGGTGTTGCT GTGTGGTCTT GGCAAGGTGT CCGCATCGAG 5640
GAGAGCGAGG CGCGCCGCAA CCGCCTGCAG GGCACAGGGG CATGCGTCCG ATCGTTTGAA 5700
GAGGTCGGCG GGGGACTGTT CATCTATCGG CCGGTGTTGC GCTGGACGGC CGAATCCATA 5760
TTCGAAGCGC ACCGCGTAGC CGGTATCAGA CCGAACCCGC TCTACCTGCA AGGCCGCAAG 5820
CGCGTCGGTT GCCTCTGCAT CAATGCTGGT AAAGACGAGA TCCGGCAATG GGATATGCGT 5880
GACCGCGACC ACATCGAGAT GATCGCCGAG TGGGAAGGCA TCGTTTCGGA TGCGTCGAAG 5940
CGTGGCAACT CCACCTTCTT CCCGGCCCCA GGCGAGACGG ACACGGCGCG GGAGCGAGGA 6000
AACATCTGGC AAGTCGTCGA GTGGTCGAAG ACGACTCGTG GCGGTCGGCA ATACGACCTT 6060
CTCGCGGACG CCGAGCCGGC GACCGCATGC GCTTCGGCGT ATGGACTTTG CGAATGAATC 6120
AGCTCCCGAA TCCTCTCACC CCAGCGGATT GCGATCTGCG CAATTTTCGC GAAATGCCGA 6180
TCGACGTGCC GCGGCTTCTC GGCTCCGATC TCGTGCACGA CGAGTCGCCC GAAGCGTGCT 6240
GGTCGGCGAT GCTGCTCTGG TGCGTCTCGT GGCATGAGGT GCCGGCAGGA AGCATGCCCG 6300
ATAACGACGA GTGGCTCGCG AAGCGTGCGG GCTACTGGCA CAAGGGCAAG CTCGATCCGA 6360
CGTGGCACGA CGTGCGCGCC GGCGCACTGC ACGGCTGGAT CAAATGCACC GACGGGCGCT 6420
TGTATCACCC GGTCCTCGCT GAAAAGGTCA ACGCCGCGTG GTTTTCGAAG CATCGCCACG 6480
CGCACGACAA GCTCGGCGAG CGTATCCGGA AGCGCAACAA GGCTAGGGCG GAACGCGGGC 6540
TGGCTCCGCT GGAAGTACCC GAGCTGGACG CGTGGATTGA CATGGGCGGC CGCTGGAACG 6600
CGAACTTTTT CCGTTGGAAT TTAGTAGTGC TTCCGGTGAA AGTAACAAGG GTTCCGGCGG 6660
AAAAGGGAAA TCTTCCGGCG GAAACGGCAC CCCTTTCCGG ACTCATGACG CTGGAATTCT 6720
AGCGGAAGGG GCCGAAAGTT CCCCCGGAAT TCCTCCGGAA AACGCTCTTA AGGGAGCGGA 6780
GCGGAACGGA ACAGAACATG TAAACCGAAG CGGCGGCGGC ACAGCACAGG CAGTAGCGCG 6840
CGAAGACGCG CCGATCGCCG CCGCCGCTTT CGTCGAAATC CTTCGCTCGT CTGGCGTCGG 6900
CTTCGCCGCC GATGACGCGC GGTTGGCGAG CTGGCCCGGG CGCGGCGTGA CGGACGCCGA 6960
CCTCCGTGCG GCCATCGGCA CGGGCCGCAA GCGCCGTGAG CGCGAAGGCA GCGGACAGCC 7020
GCTGAACCTT GGCCTGCTCG ACATGATCCT CGGCGATCTG CTTGCCGCGC GCGCCGCGAA 7080
GCCGGCCACG GGGACGCGCG CCGTCGGCGA CTGGTGGCGC TCGTGGACCG GCATCGTCGA 7140
GCATGGCGGC ACGCTCGGCG TCGAGCAAGG GCGCGACGAA CCTCCGTTCG ATTTCAAGCT 7200
GCGCGTCTTC GACGCGGCCG GCGACGGCCC GTGGTGGGAC GAGCACAACC GCGGATTTCG 7260
CAACACCGGG CCAGTATCCG CTGGCTCGTT TCTGGAGAAC TCGCAATGCT GACGGGGAAC 7320
ACTCGCTATC GGCTCGGCTG GTTCGGGAAG GTGATTCTCG AGGTGTCGGA ATGGCGACGC 7380
CGACCAACCC GCGACCACCC AGGATACCTG CCCGTGGATC GAGGATTGGC GCGATGCCAC 7440

GTTTCGCGAC GTGCTTGACC TGGCGGAGCG GAATTGGGCA CTCGTGCCGC GCCCTTCGAA 7500
CGACGACCGG AGGGGAGCGC GATGAACTGC AAACCCGGCG ACATGGCCGT GATCACGCAT 7560
GGCCTCGCAC GCGATCGAAT CGTTGAGGTG AAGGCGCCGT ACGGCGATTA CATGCATTTC 7620
GGCTTCTGCT GGTACGTCGA AGCGCCGGCA CCAATCCCTG CCATGGACGC GAGAACTGGC 7680
TTGTCGTGTC AGATCACGAA CGGCTGGTGC CCGGATGCGT GGCTGCGCCC GATCAGCGGC 7740
GTGCCGGTGA CCGACGACGT CGAAGGCGAG GTGACGGCAT GAGGTCTTGG GAATGGATGT 7800
GATCGAGTCG CCGCTGTGCG CTGACGTGCC GCGCATGACC GTTTCGCCGC GATTCGCCGA 7860
GCTGATGCCC GCCGAGTTCG TGAGCGACTT GAACCGGTGG ATGCGCGAGT TCTTCGGCAC 7920
TTCCGATGTG TCGTACATCG TCGAAGGCCG CATGATCGTC ATGAGCCCGC GCGCGGCCGC 7980
CGATCTTCGC CGCCAGGCTT CGCAGGCTTC GGTTCGTCCC GGGCATGGTC TAGGGTGGCG 8040
GCCATGACGC AGCAATCTCT CATCACGGCA TCGCCGATCG CGCAGCGTGT CGAGTTCGTC 8100
GTTCCCGGCA CGCCGGTCGC CAAGGGGCGC CCGAAGTTCG CGCGCCGCGG CGAGAGGGTC 8160
ACGACCTACA CGCCCGAGAA GACGGAGCGG TACGAGAACC TCGTGAAGAT GTCCGCGCGA 8220
GCGGCGATGC GCGGCAGCGA GCCGTACGCC GGCCCGATCC GGCTGATCGT GCACATCGGG 8280
CTGCCGATCC CGGCGAGCTG GTCGATGAAG CGCCAGGGCG AAGCGGCAGT CGGCGCCATC 8340
GGCGCGACGA AGAAGCCGGA CGCCGACAAC GTTGTCAAGG CGTTGAAGGA CGGCATGAAC 8400
GGGGTGGTGT ACGTCGACGA CGGCCAGGTC GTCGACCTCT GGGTGTCGAA GCGCTACGCG 8460
CGCACGCCGG GTGTGCGCAT CGAGGCGATC GAATTGAATT TGAAGTCAGC ATAGGGAGCG 8520
GGGCCTTGAA AGCTAAAAGC AAACTCACCA TCAACGAGAT GATCGGGAAG ATGAAGCCCG 8580
GCGTCCGGTA CTCGGCGCAC GACCTCGCCC GGCGCCTGAA GCATCCGGTT TCGTCGGTGC 8640
GCCAGCTGCT CGCGCTTGAC GTCGCGCTCG CACGGCTCGA CTGCCACTCG GAGAACCGCG 8700
GTCGGATGTA CTCGCTCGCG GGCACCAGCC GCTCACCCGG GTCACACGTC GACACGCGCA 8760
TCCGCCCCGA CTTCACCAGC AACCTGACGG GATACATGCA CGAGCTCAAC ACGCGCAAGG 8820
CGCTGGCGAT GACGACGCGG GGTGCACGAT GAGCGACGTC GTCGAGTTCA AATCGGCGTT 8880
CGACGCCGTG CGGTTCGCGC TCTGCTACTC GTCGCAGCAG TACGGCGAGA CGATGCTGGC 8940
AAAGCGGCTC CGCGGCGAGT CGATCGGCAC GGGCATGGGC CTAGTCGGAC TGGACGGCGC 9000
CGGGCAGGCC GGGGAGATCC GCCGGCATCT CTGGGACCTG CCCGAGCTGC ACCTGTCCGT 9060
GATCGTCGCG CGCGCGGCGC CGCGCGACCT GCCGTGCTCA TGCGGCGCGG CGTGCTGCAG 9120
CGGCCGGACG CCGAATCTCG AGTGGAAGGC GGCGATCGGG TGGCTGACGC AGGCGTCTGC 9180
CGCCTACTGC TCCGGCTTCT CGCACTTCCG CGTGCGGCGG GCAATCATCG AGCGGCTGTT 9240
TGGAGTGAAG TGCGACCTGG TCGAAATCGC CCAGGACTGC GACGCGCACG TGAACACGGT 9300
CAGCAAGCAG AATGCCGCGG TGCGGAAGTG GATCGAGGGG GACAAGAAGA GCGGTGAGGC 9360
GGGCGTCGAG GGGCTAGCGT GGTCGGTGAT CGAGCGACGG TTCAGCGAAC TCGGACTTGC 9420
TGCAGGAACG AACGACTGCT TGACAATGGT GTGTTTCGCA CACAATAATC CGCCATATTC 9480
GATACACGTC ATACGTGCGT CCGAAGCAAA AAAGCCCGCG AAAGCGGGCT TTTTTGTTGC 9540
CTCGATCAGG CCTTGTTGAG ATAGAAGTAC GACGATCCTT CCGTGGTAAT TACGGAGTTT 9600
GTTTTGCCGG TCGCGTTGGC TGCTTCCCGC GCATGCTGAG CGACTTGCTC GGTCGTGAGC 9660
GTGCTTTCGA CGTAATAGGT GGCAGTCGGT AGCTCGTAAT GGTCGCCGCT GCTTGACTCG 9720
ATTGTTCGTA GAAAGCCTTT AGCTTCCATT CGCGCGTGAA GGTCGTTGTA ATCATCCCAA 9780
CTCGCCGCAT GAAGCTCAAT GCGCGTGACG AATCTTGCCA TTTGTGTCAC CCCGGTGGAT 9840
GTGAGCCGGT AAGGGCCGGC TCGACGATTT TACGCCGATC GCCTGATCGT CCCGAGCACG 9900
TGAAAAACAT GCCTCGTCCG AAGCGAACTG ACGTCGCGCA CGTCGCGCAT GCGCCGCGGC 9960
CGGCTCCCCC CGAAATACTC TTTGGCGATT CGAACTGGAT CCGGCGCATC GTACCGGCCG 10020
ACGGCATCTC CGAATGGGTG AACGAAACTC TCCTGCGCGA TGGCGCGCCG CTGCACAACC 10080
CTGACCACGC GCACCTAGTC GACGCCGACG TTGCCTACCT CTGGGCGGCT GTCGAGAACG 10140
TGCGCCAGAT GCGCCGCGTC GTCGGCCAGT GCGAAGAGGT GATGATCCGC GCCGGCGGCT 10200
GGCAGCGCGC GCGGCAGGAG CAGCAGTTCT GCGAATGGTT CGGCCGCGTG CCGGCCTTCC 10260
TGATCACGCT CGACGCGCAC TACGCGCGCG AGTGCAGCGA CCTGCAGTGG TGCGCGCTGA 10320
TCGAACACGA GCTATATCAC GTGGGTCAGC GGCTCGACGA GTTCGGCGCG CCGGCGTTCA 10380
CCAAGGGCGG CCTGCCGAAG CTCGGCATCC GCGGGCACGA CGTCGAGGAG TTCGTCGGCA 10440
TCGTCCGGCG CTACGGCGTG GCCGGCGGCG CGGGCGACAC CGCGAAGCTC GTCGCCGCGG 10500
CTCAGCGCGC GCCCGAGGTC GGCCATGCCG ACATCGCGCG CGCCTGCGGC ACCTGCATCC 10560
TGCGGGCCGC CTAACCCGAA CGTTTTCCCG CTATGGCAGC ACTTCCCGAC GCGATCAAGG 10620
TGTTCATCGT GCACGTCGCT GGCGTGCTTC GACACGATCT CGCGCACAGC GAAGGCCGTG 10680
CGCGAGGAGT TCGGCGTCGA GGTGTCGCCG CAGCAGTGCG AGCGCTACGA CCCGACGAAG 10740
CGGGCTGGCG AGACGCTCAG CAAGAAGTAC CGCGAGATCT TCGAGCGCAC GCGCGAGGAG 10800
TTCCTGAACG ACACGTCGCG CATCGGCATC GCGCACCGCG CGGTCCGCCT GCGCAAGCTG 10860
GCGATGGCTG TCGATAAGCC CGAGGAGCGC GGCAACCTGG TGCTGATGGC GCAGTTGCTT 10920
GAGCAGGCGG CGAAGGAGGC CGGCGACGCG TTCACAAACC GTCATCGGCT CGAACACACC 10980

GGCAAGGGCG GCGGCCCGAT CACGGCGATC TCGACGCAGA CGAACGATCC GATGGAGGCG 11040
GCCAAGATCT ACGCCGAGCT TATGAAGCCA TAGCATGCCC ATCCCGTTCC CGTTTGACTT 11100
CCGCGCACCG GACTATGTGC AGGTGTTCGA ATGGCGAGCG GAGCGCCTGC AGCGCATCCG 11160
CGCGAACCCG GGCGTGCTGC CGGCGCTGCG GGCGTTCTAT CGCGACAACC CGGCCCAGTT 11220
CATCATCGAC TGGGGCATGA CGTTCGATCC GCGGAACGTT GAGCGCGGGC TGCCGGCCAC 11280
GATCCCGTTC CTACTGTTCC CTAAGCAGGA AGACTGGATC GCGTGGTTCA TGGAGCGCTG 11340
GCGCGCGCGC GAGCCCGGCA TCACCGAGAA GACGCGCGAC ATGGGGATGT CGTGGCTGAC 11400
GGTCGGCCTC GCCGACACGG TGTGCCTGTT CCACGAGGGC GTCGCGGCCG GCTTTGGCTC 11460
GCGCAAGGAA GAGTACGTCG ACAAGATCGG CTCGCCCAAG AGCCTGTTCT GGAAGGCGCG 11520
CGAGTTCCTG CGGCTGCTGC CGGCCGAGTT CCGTGGTTCG TGGGACATCG GCACGCACGC 11580
CCCGCACATG CGCATCATCT TCCCGGACAC GGGGTCGGTG ATCACCGGCG AGTCGGGCGA 11640
CGGGATCGGG CGCGGCGACC GCGCCAGCTT CTATGTGGTC GACGAGTCGG CGTTCCTCGA 11700
GCGACCGCAG TTGGTCGACG CATCGCTGTC GGCCACGACG AACTGCCGGC AGGACATCTC 11760
GACGCCGAAC GGCATGGGCA ACTCGTTCGC GCAGCGCCGG CACAGCGGCA AGATCAAGGT 11820
GTTCACGTTC CACTGGCGCG ACGACCCGCG CAAGGACGAC GCCTGGTACG CGAAGCAGGT 11880
GGCCGAGCTG GACCCGGTCG TCGTTGCGCA GGAGATCGAC ATCAACTACG CGGCGTCTGT 11940
CGAGGGGGTC GTGATCCCGT CTGCCTGGGT GCAGGCAGCG CTCGGCGCGC ACGTCAAGCT 12000
CGGCATCGAG CCGAGCGGCA CGCGCCGCGG CGGCCTCGAC GTCGCGGACG AGGGCAAGGA 12060
CAAGAACGCA TTCGCCGGCC GCTACGGCTT CCTGCTCGAG CATCTCGAAT CGTGGTCCGG 12120
CGTGGGCGGT GACATCTTCG GCACGGTCGA CCGCGTGCTC GGTATCTGCG ATGTGCGCGA 12180
CTACGAGGTG TTCGACTACG ACGCTGACGG CCTTGGCGCC GGCGTGCGCG GTGACGCGCG 12240
CGTGCTGAAT GAGCAGCGCG TGGCGGCCGG CAAGCGGTCG ATCCGCAATG AGCCGTTCCG 12300
CGGCTCCGGG CCAGTGTACG ACCCGGAGGG GAGATGGTGA AGGAGCGGAA GAACAAGGAC 12360
TACTTCGCGA ACCTGAAGGC GCAGTCGTGG TGGGCGCTGC GCCTGCGCTT CCAGGCCACG 12420
TATCGCGCGG TAGTGGAAGG AAAGCCGTTC GATCCGGACG AGATCATCTC GATCGATCCG 12480
GATCTGCCCG AGCGCGCAGC GCTGAGCATG GAACTGTCGC AGCCGACCTT CACGGTCAAC 12540
GGTGTCGGCA AAATCGTGAT CGACAAGGCG CCCGACGGAA CGAAGTCGCC GAACCTCGCG 12600
GACGCGGTGA TGATTGCCTA TCAGCCCGCT GTGCGCGGCA TTGATATCTG GAAGAGGCTG 12660
GCAGGATGAG TCGAAAGAAC CGAAGCGGCG TGCGCGGCGC GTCCGTGACC CGTCCGTCGT 12720
CGGCCGGCTC GTCGGCCAGC ATCCGCACGG CCGACTCGTT CGCCAACTTC GAAGCGCGGC 12780
TCGGATGGGG AGCCGACAAC CAGGCGTCGG CAGCGCAGTA CACGCTGTCG TACCAGAGCC 12840
GCAACCGCGT TTGGCTGGAA GCGGCTTATC GCGGTTCGTG GATCGTGCGC GCCGCGGTGG 12900
ACGCGATCCC GGAGGACATG ACCCGCAAGG GCATCGAGAT GTCCGGGCTC GATCCGACCG 12960
ACGTGTCCAA GATGGAGACG GCGCTGACGC GCAAGGCGAT CTGGGACCAG CTCTGCGACA 13020
CCGGCAAGTG GGCGCAGCTG TACGGCGGCG CGATCGCGGT GATGCTGATC GACGGCCATG 13080
ACATGTCGCA GCCGCTTCGG CGCGAGACCA TCGGGAAAGG CCAGTTCAAG GGCCTGCTCG 13140
TGCTCGACCG CTGGATGGTT GCGCCGCCGG TCGGCGAAGT GGTGACCGAG TTCGGCCCCG 13200
ACCTCGGCAT GCCGAAGTAC TACGACGTGC TTCCGACGAC GATCGGATTG CCGCAGGGGC 13260
GCATTCACCA CTCGCGGGTG CTGCGCATGG ACGGCGAAGC ACTGCCGTAC TACCAGCGCA 13320
TCAGCGAGAA CGGTTGGGGC CTATCGATCC TCGAGCCGAT GTGGGACCGG CTTATCGCGT 13380
TCGATAGCGC GACGGTCGGC GCCGGGCAGC TCGTCTACAA GGCGCATCTG CGCACGCTGA 13440
GCGTCGAGAA GCTGCGCGAG ATCATCGCGG CGGGCGGCCC GGCGCTCAAC GGCCTGCTGA 13500
AGCAGGTCGA GATGATCCGG CTCGGGCAGT CGAACGAGGG CCTCACCCTC ATCGACGCGA 13560
CCGACAAGTT CGAGACTCAC CAGTACGCGT TCAGTGGCCT GTCCGACGTC TTGCTCCAGT 13620
TCGCGATGCA GCTCAGCGGC GCGACGGGCA TTCCGCTCGA TCGCCTGTTT GGCCAGCAGC 13680
CGGCCGGCCT GAGCGACACC GGCGAAGGTT CGCGCCTGCT GTATCACGAG AAGGTGCACA 13740
CGCGGCAGGA GCGGCGGCTG CGCAATCCGT TGCACGGGCT GCTCGATGTG ATGTGTCGTT 13800
CGGAAATCGG TCAGCCGCTC CCCGAGGACT TCTCGTACGA GTCAACCCGC TGCGGCGCCA 13860
GTCCA 13865

Figure A.8.3 Contig E 2306bp

GGCGATGCGT GCGCTCCTGT CCGAGCAGGT GCGGCTGATC AAGCCAATTC CGCTGGACGC 60
AGCCGAGCGT GTGCACCGGC TGACGCTGGA AGGAATCGTC GACGGCACGC GCGCCGCGCA 120
GATCTCGACG GCGATTCAGG AGTCCGGGCA GGTCGCGAAA AGCCGGGCCG ACACGATCGC 180
GAGAACCGAG GTCAGTCGCA CGGCCGCGAC GCTCACCGAG GCGCGCGCGC TCGACGTGGG 240
CAGTCCCGGT TACTTCTGGC GGACGTCGGG AGACTCGGAC GTGCGCGAGG ACCATCGCGA 300
GCTGGAAGGC AAGTTTTTCA CGTGGGACAA GCCGCCGGTC GCGGATAAGC GGTCCGGCGC 360

```
GCGGGCGCAT CCGGGCTGCA TCTACAACTG CCGGTGCTGG GCCGAAGTCG TGCTTCCGAA 420
GGGGTGATAT GGCGAAGATG AAACGCGACA GCAGCGGATC GTTGCTGTTC GAGTGCCCAT 480
GCGGCGAGCT GCACGTCGTC TATCCGCATG GTTGCGATTC GCAGAACCCA GCGCGATGGA 540
ACTGGAACGG CAGCGTGGAT GCACCGACGT TGTCTCCGTC GATCCTCGTT TCGTGGCCCG 600
GGCCCGGCGA CCAGCAGAAC GTCTGCCATT CATTCATCAC GGATGGCTGG ATCCAGTTCT 660
GCGGCGACTG CACACACGAA CTCGCCGGGC AGACGGTCGA GATTCCCGAC TGGGTGGACT 720
GAACGTGACG GCACATCGAA CGACGTTGCG CGTGCGAGTG CGCACCGCGT GGTGGTTGCC 780
GCTATACGTC CGATCCCTGG CCGTCTGGTG CCGGCTCACG CGCACCGAAC CGGACTACGA 840
GCGAATCCGC GGCGTCATTG CGCGCGGCGT CCGATCCAGC ATCGAACGCT GAACGTTCAG 900
CAATCGCGTG CACTCCGTAG CAGCCGGGCC AAGCCCGTTA GACCCGACTG ATCAATGCCC 960
AAGCATCTGC ATATCTACTT CCATACGTAT GACGCGTCGT GGGAGGAATC GAAGCATCCG 1020
CGAGCCGCGA ACGGCCAGTT CGGAGCAGGC GGGGCCACTG GGCCGGCGGC GGCCGCGCCG 1080
ACGACGCTGA AGGGCAATGA GCTTGGCGAC TTCACCAGCA TGAAGGAACT ACGCCAGAAG 1140
GCGATCGCCT ACGGCAAGCA GTTCGCTGGG AAGAAGTTCA AGAATGCGGC GACAGGCAAC 1200
CAGATCGAAG TGACGAACGG CGGGATCCGG CACACCGTCG CGACCGGCCA CGATGAGGTG 1260
CTGCGCTCAA TCCCGGCTCT GCCCGACCTG TTGACGAAGG CTCGGCTGAT CGACTCGCAG 1320
CCCGACAAGC GCGGCGACCC GAACGTCAAG GCGGTCGAGA CGTACTCAGC TCCGCTCAAG 1380
CTGGACGGCA AGAACTATCG GGCGGTGATC ACGGTGAAGG TGTTCTACGA CGGCCATCGG 1440
TATTACAACC AGGGCCTGGT GCGGGAAGGG GAATAGGGCC GGCCGTCGTT TAAATAAGGC 1500
ACCCCCTGCC TTTCGGCAGT CGGCTCACCT CCGGCGACCG TCCCAGTACG CATTATAGCC 1560
AAGTCCTCAC GACTTGAGAA TTTCCTGTCG AATCACGAAC GTACATCCAA TTTCGACCCA 1620
GATGCGGACC ATTCGAATTC CTACGGCAGA CCACGCATGC AGTTGCGGTG CCGGTGCCGC 1680
GCGCGCTCGC GCGCATACCC GCGATGGCAT CACCGCGTCG GGCGTGTACG CGGCCGAGCA 1740
GCTCGGCGAG CGGCAGTCGA TCACGCCGGA AGGTTTCCTG CTCTGCGAGG CCGTGCCGAT 1800
CGCGCGCGTC GGCGCGCAGG ACTATGCCTA CTTCGAGTTG CCGGAGATCG AGGCGAAGGA 1860
CGGCGTTATC GTCGCCGAGC GCACGGCCGA CGTGCTGTTC AGCCCCGAGA CGCTCGCTAG 1920
CTTTGAAGGC AAGCCGATCA CGATCGACCA TCCGCCGGAC TTCGTGACGC CGGCGAACTT 1980
CAAATCGGTC GTCGTCGGTT TCGTGATGAA CGTTCGACGC GGCGAAGGCG ACCAGTCGGA 2040
TTTGATGCTG GCCGATCTGC TGATATTCGA TGCAGAAGCC ATTCGCCTCG TGCAGCTCAA 2100
GGTTCTGGCA CAGGTCAGCA ACGGCTACGA CGCCGACTAC GAACAGATTG CGCCTGGGCG 2160
GGCGCGACAG GTGGTGATCG TGGGCAACCA CGTCGCCCTC GTGAAAAGCG CCCGCTGTGG 2220
CCCCGTGTGT TCGATCGGGG ATGGCAGTTC CAACCTACTC CCGACAGGAG ATGCAAGCAT 2280
GGCAACCAAG AAAGGCTCCA AGTTCG                                     2306
```

Figure A.8.4 Contig R 227bp

```
AACGAGAACG CCGAAATCGT CAAGTCGGTG GTCGGCGGCG CGAACGTGGC CAGCATGACC 60
TGCGATTCCG TCGCGGCGTT CTTCAACGCA GCGTCGGAGG TCGTGCGCAG CAAGAACTCC 120
GGCGTCACGC AGCGCCGGAC GAACGATTCC ACTCAGACCG AGCGGAAAGA CATCAACGAA 180
ATCCACGCGG AATTCTGGAA GGTCCGCAAG TAAGGAGCCG ACATGCC              227
```

Figure A.8.5 Contig Q 376bp

```
CGGCATGCAG CTTGCATGCC TGCAGGTCGT TTACCTTACC GAAATCGGTA CGGATACCGC 60
GAAAGAGCAG ATTTATAACC GCTTCACACT GACGCCGGAA GGGGATGAAC CGCTTCCCGG 120
TGCCGTTCAC TTCCCGAATA ACCCGGATAT TTTTGATCTG ACCGAAGCGC AGCAGCTGAC 180
TGCTGAAGAG CAGGTCGAAA AATGGGTGGA TGGCAGGAAA AAAATACTGT GGGACAGCAA 240
AAAGCGACGC AATGAGGCAC TCGACTGCTT CGTTTATGCG CTGGCGGCGC TGCGCATCAG 300
TATTTCCCGC TGGCAGCTGG ATCTCAGTGC GCTGCTGGAC TCTAGAGGAT CGCTGCCAGC 360
GAGCGCTGGA GCTCGA                                               376
```

Figure A.8.6 Contig O 435bp

```
GGGTGCAGCG GGGCGACCTG CAGACGCCGA ATTCGGCGCG CATCCGGGTC TACAACGTGT 60
CCGCGACGAC AGCGCGGCGC GCGCGGAAGG AGTTCACGCG CGTCGTGCTG CAGGCCGGCT 120
ACGAGGGCAA CTACGGGATC ATCTTCGACG GTCAGATCAA GCAGGTGCGC CGCGGGCGCG 180
AGAGCCAGAC CGATACGTTC CTCGACATCA CGGCGGCGGA CGGCGACTCC GCGTACAACT 240
```

TCGCGGTGGT GAACACGACG CTTGCGGCCG GCTCGACGCC AGCTGATCAC GTAGCTGCCG 300
CCTGTACGGC AATGAATCCG TACGGCGTGC AGCCGGGCTA TCTCCCTGAG CTGCCTTCGA 360
ATCCATTGCC CCGCGGCAAG GTGATGTTCG GCATGGCGCG CGATTTCATG CGCTGGAACA 420
CGCGCAGCCA CGGGC 435

Figure A.8.7 Contig DU 4870bp

TCGACGATCC GTTATCTGGC GATGGACGAG TTCGCGCTCC ATAAAGGCCA TCGCTACGCC 60
ACGGTGGTGG TTGATCCGAT CGGCCGGCAG GTCCTCTGGG TTGGGCCCGG ACGGTCACGC 120
GAGACGGCGC GCGCCTTCTT CGAACAACTC CCCGAAGGCG TGGCCGAGCG CATCGAAGCG 180
GTCGCAATCG ACATGACCAC GGCCTATGAG CTGGAGATCA AGGAACAGTG CCCACAGGCG 240
GAAATCGTCT TTGACCTGTA CCACGTCGTG GCCAAGTACG GTCGCGAGGT GATCGATCGG 300
GTACGGGTGG ATCAGGCCAA CCAACTGCGA CATGACAAGC CGGCCCGCAA GGTTCTGAAG 360
TCCAGTCGCT GGTTGCTGCT GCGCAACCGT CATAACCTGA AGCCAGAACA GGCCGTGCAT 420
CTGAAGGAAC TGCTGGCGGC CAATCAGTCG CTGTTATGCG TCTATGTGCT GCGCGACGAG 480
CTCAAACGGC TCTGGTTCTA CCGCAAGCCG GCCTGGGCGG AAAAGGCTTG GGGGCAATGG 540
TTCGAACAGG CTCAGCAAAG CGGGATCGCC GCCTTGCAAA AGTTCGCCCA GCGCTTGCAG 600
GGTTACTGGC ACGGAATCGT GGCCCGCTGC CGCCATCCGC TCAATACCAG CGTCGTCGAA 660
GGCATCAACA ACACGATCAA GTCATCAAGC GCCGCGCTTA CGGGTACCGC GACGAGCAAT 720
ACTTCTTCCT CAAGATCCGC GCCGCGTTCC CCGGGATTCC GCGATGAACC AGATTAAAGC 780
CCGTTATGGC GAGCAGGTCC TATCCCTGCG AATTCCTTGA CTACACCCCG CTCGCGCGGG 840
GTTTTTCGTT TTTAGGGACT TATGCGGTAC CGAAAACTCG ACGCCGACGG CGACTATGTC 900
TTCGGCGGCG GCGCGGCCGA CTTCCTTGTG AACACGCCTG AGACGGTTGC ACAAGCCGTG 960
TTGACGCGCC TGCGCCTGCT GCGCGGCGAA TGGTTCCTCG ACACGACGGC CGGCATGCCA 1020
TGGGCGACCG ACGTGCTCGG AAAGTACACG AGCGGCAAGT ACGACGCGGC GATCCGCCAG 1080
TGCATCCTCG GCACGCAAAG CGTGACCGAG CTCGTTAGCT ACTCGAGCAC GGTTGATCCG 1140
GAGACGCGCG TGCTGACCGT CACAGCGACG ATCAACACCA TTTACGGCAC CACCACGGTA 1200
CAGGCGACAT TGTGACTCTC ACGACCCTCG CACCCACCAT CGACGCGAAC GGCATCACCG 1260
CGCCGACGTA CGCGGACGTG ATCGCGTATC TGCAGGACCA GTACCGCTCG ATTTATGGCG 1320
CCGACACGTA CCTGGAGTCG GACAGTCAGG ACGGCCAACT GCTCGGCGTG TTCGCGAAAG 1380
CGATCAGCGA CGTCAATTCG GTCGCGGTCG CGATCTACCG GTCGTTCAGC CCGGCCACCG 1440
CGCAGAAAGA CGCGCTGTCG AGCAACGTCA AGATCAACGG CATTGCGCGC AAGGTCGCGT 1500
CGTATTCGAG CGCCGATCTG GTGCTGGTCG GGCAGGCTGG CAAAACAATT ACGAACGGTG 1560
CTGCGAAGGA CGCCAACGGC GTACAGTGGA TGCTGCCGGC TAGCGTGACG ATTCCGCCGA 1620
GCGGCACGAT CACCGTTACG GCCACGTGCG CGACGATCGG CGACATTTCT GCGCGCGCTG 1680
GCGCGATCAA CCAGATCGCG ACGCCGACAC TCGGCTGGCA GTCGGTGACG AATCCGGCGG 1740
ACGCCGCTGA GGGCGCGCCA GTCGAGAAGG ACGCGGTGCT GAGGCAGCGG CAGACGGTGT 1800
CCACCGCGCT ACCGTCGCTC ACGGTGCTCG ACGGCATCAT CGGCGCGGTG GCGAACGTTC 1860
CGGGCGTCAC CCGGTATGTC GCATACGAAA ATGACACGGA TACGACGGAC GCGAACGGCA 1920
TCCCGTCGCA TTCGATTTCG CTGGTCGTCG AGGGCGGCGA TGCCACGGCG ATCGCAAATG 1980
CGATCGCGGC GAAGAAGACG CCGGGTTCCG GGACGTATGG CACGACTGCC ATCATCGTCG 2040
CGGATATCTA CGGCCGTCCG ATCACGATCC GGTTCTTCCG ACCGGTGGCC GCGCCGATCG 2100
GCGCCACGGT CACGATCAAG GCGCTCACCG GCTACACCAG CCAGGCGGGC CAGCAGATTC 2160
AGCAGGCTGT GTCGGACTAC ATCAACGGCG TGCAGATCGG CGGCGGCCTA TCTGGCAGCG 2220
TCGAATGGGG TGACGCCCTG ACCGCGGCGA ACAGCGTCGG TGGTGGGGTG ACGTTCAAGC 2280
TGTCGGGCCT GACGCTGACC GGGCCGCGCG GCGCCGGCGC GCCGGACGTC GCGCTGTTGT 2340
TCAACGAGGC GGCGTCTTGC ACGCCAGCGA ACGTGACACT GGTGGTGACC TGATGGCGGA 2400
TCTGACCGAT TACACCGTGT TGATCACATC GGAGCACAGC GACAAACCGC GCTTCATGGC 2460
AAGCGTCAGC GCGCTTGTGC AGCCGCTCGT CGAACAGATG AATGTGCTGG AGAGCATGCC 2520
GGGTAAGTTC GACCTCGACA ACGCGGTCGG CGTGCAACTG GACGATGTCG GCCTCTGGGT 2580
TGGCGTGTCT CGGAAAATTC GCACACCGTT GACCGGCGTC TACTTCTCGT TCGACATCGC 2640
GGGCCTCGGC TTCGATCAAG GCACGTGGAA AGGACCGTTC GATCCCGATA CGGGGCTCAC 2700
GATCCTCGAC GATGACACGT ATCGATTGGT CATCCGCGCG AAGATCGGCG CGAACCACTG 2760
GGACGGGACG CTGCAGCAAA GCGCCGCGAT CCTGAACAGC ATCTTCGACG CGGATACGCA 2820
CGTCTTCATC GAAGACCACC AGGACATGTC GATGACGATC GGCATTGCCG GGAAGGTCCC 2880
GCCGGCGACG TTCCTTGCGC TTTTGTCAGG GGGCTACATC CCGCTCAAGC CTGAAGGCGT 2940
CCGCGTCAAC TACACGATCG TGACGACCGT CGACGGATCA CCCCTGTTCG GATTCGACAT 3000
GAGCAATCAA CTCGTGGCCG GATTCGACGT TGGGGCCTGG AGCCGCCCCG TTTAACCGCC 3060

AATTGCATTG TCTGCAAGCC ACCTTCGGGT GGCTTTTTTT ATGCTCGGAG CATTGATGGC 3120
AACGAACGAC TTTCTCGTGT TCGGCGGAGG CAGCTCCCCG AACGTTATCG ATCAGGCGAC 3180
CTACGCGGCT CTCACGGCCC GTCTGTCGGG ATTTCAATCT GGCGCCGCGC TGTCGGCGCA 3240
GCTCAACAAG GTATGGCGCC AGAGCTCGAT CATGGCGGCG GTACTCGCGC AGTTCACGGC 3300
GAACTTCTCG GGCCAGAACT CCGTCGACGA CGGCACGATC GCGACGCTGT TGGCGAACCT 3360
ACAGGCAGCG ATCAACGCAG CAGGGATCAC GGCTCCGCAG TTCGACAACA GCACGAAGCT 3420
CGCGACGACG GCATTCGCGC AGCGGTCGCT TGGCAATTTC CAAGCATTCT ATGCATACAC 3480
GTCGAGTCAG ACTCTCACGG CTTCGCAATC TGGCTCAGTC ATCAATTTTT GGGGAAATTC 3540
CGCGTCGACT TTCACGCTTC CATCGTGCTC TGCAATGCCT GCTGGCGGAT CGTTCTTGTT 3600
CAACAATTCC AATGCCGGCA ACACTTCGGT CACGGTGACT CGGGCAGGCT CAGATTCAAT 3660
TCTTTGCGGA GGCAATGCCA CTAGTGTCGT TGTCGGGCCG GGCGATAACC TTCTCTTGGT 3720
CGCGATTCCG CCGAGCCAAT GGGTTGCGAC TGGCGGAAGC GCTCAATTGC CATTTGCGAG 3780
TACCGCGCAG CGGACATCCG GCGGCGTCGT CGGAGCCATG CGCAACGCCA GCATGAACGT 3840
AGCAGCAGCG AGCGCATCGG CGACCTTCAC TGCTGATGAG ATCGTCGTCG AAACCGCGCT 3900
CGGTGGCGGC TTCTACCGGC TCGCCAGTTT CAGCAAGACG ATCAACCTCG CGACGACCGG 3960
CGCGGGCGGC ATGGACACGG GCAGCGCGCC GGTGAGCGGC TACGTTGCGC TGTACGCGAT 4020
CTACAACCCT ACGACCGGAG CAAGCGCACT GCTCGCGAGG AACGCAACGA GCGCGGTGCA 4080
GGGGAGCGTG TACGGCGGCG CGAACATGCC GACCGGCTAT ACCGCTTCGG CGCTGGTGAG 4140
CGTGTGGCCG ACGAATGGGA GCGAGCAGTT CATCACGGGC GTGCAGGTCG ATCGAGTGAT 4200
CTCGATCCCG TTCGTTACTG TGCTGACGTC GAGCACCACG CAAGCATCGC CGACGGCACT 4260
GTCGATTTCC TCCGCAGTGC CGCCGAATGC GCGAAGGGTG TCCGGTACGA TATCAGTATC 4320
GTCGACGTCC TCGACGCCGA ACTCGTCGCT CAGTGTCTAT GCGGCATCGA CGAGTGTCGG 4380
CATTCAGGGG TTGAACAATA GCGTGAGCGC TTCGGGCGGC ATCTCGACGA ACTACAAGGA 4440
CCTGCCGATC ACTGTTTCGC AGACGCTCTA TTACACAGCG ACATCATCGG CCGGTACGCC 4500
GACGTTCACG ATAAGCGTTG CCAGTTATGA TTTTTGAGGA GCTGACATGT TCGTACAGTT 4560
TTCTGACGCA GACGAGAAAG TCATCACCGC CGTATTTGCC AACGAGCAAG ACCGTGAGGT 4620
CTTCCCCAAT CAAGGGAAGA TCGATCTGAC CGATCCTCGT TATGCTGCAT TCTTCAACGC 4680
ACTGCCTCGG TTGGCTCAAC AGGCGATTCC CTCGCCGGTT GCCGACTAAA CTCCGCTCCC 4740
TTTACCTTCA GGGCTCGTTT CTCGACCAGA TGCCAGGAGG CAAAGGCGAG CGGAATAATG 4800
ATTGCCAGAG ACAGAACGAT CGAAAGACCT ACCGGGATCT CCCCTGCGAG ACGCCGCGCG 4860
ATGATTTGTT 4870

Figure A.8.8 Contig V 2618bp

CGCGCACATC GCCGCAGCGA TCAGCCACTT CTCAAGCCTG AAGCTGCTCG TGCTCGATCG 60
CGCCGACGTC CTGGTCGGCC CGGAGCGCGA TCGACTGCTC TACTGGCTCG ACGATCTGGC 120
CTACGCCGGT CAGATCGACA CGGCGCTCGT GTTCATGAGC CTGAAAACGC CGCCCGGTGG 180
CCTGCCGGAG GCCATCGAGG CTTTCTGGGT CGAGGGCGGT CAGGTCGCGC CGGCCGCGCA 240
ACATGCAATA CGGGAGGCAG CGTGAGAGAG GACATCGAGA AATATCTCGC CGCGACGTCG 300
AAAGCCACGG CGAAGGCCGT GGCGACCGGA ACCGGGCTCC CGCAGCTCGA CGTGACGAAG 360
GAGCTGAACC GGATGCTCGG CGACGGGCTG GTCGAGCGCG AGAAGCGGGC CGGGGGCGGC 420
AACGAGTACG TGTACTGGCT CGCGCGTGCA GTACAGGCGG CCGCGCCGTG CGCCGACACG 480
CCGCCGACTG CTGCCGCGCC GGCGCTGGTA TCGGTCGGCC TGGTCGAGAA GTCGTTGGAC 540
CCGAACGCCG GCGTCATCGA CGTCGCGCGG ATCATCGCGG ACCTGCGCGC CGATGTCGAG 600
CGCCTCACCG CCGAGCGCGA CGCCGCGCAG CAGCGGGCCG ACACCTGGCG CGCGAACTGC 660
CGCGTGCACT CGAAGCGCGC ATCGACGAGC TGACGCTCGG GCCGGTCGGC GCGCGTGCGC 720
CGCTGTTCGT GACGGTCGGT AGGTACTGCA AGCCGAGGCG TCACGCGTCG CTCGAGAAGG 780
CCCAGCGGCG CGGCAGCGCG CTCGTGCGCA GCGAGAAGGA ATCCGAGTGT GCTCGTGCTT 840
GAGCCAGTCG GTCGGATCGT GCGCGGAACG CAGTGGATGC CCCGATAGCA GTACCGCCGC 900
GCGCCTTCCG TGCCTCGGAT TGCGCGGCGC ATTCGGGCGG CTCGCACAGC GCCCGTTTTT 960
TAGATCTGAC CATGCAAACT GACGAAGACA CCGTGGAATC CATCACGGCG CGCGACCGCT 1020
CGCGTCAAGA ATCCTCTGCC GGCTCCGACG ACATGCCCAT ACTGACGGTG GCCCCGTCGA 1080
GATCGTGAAC AACTCGGCGA TCTACGGCCG CGAGTACGGC GAGTGGCCGT GGGCGTTCCT 1140
CTGCCGCTCC TGCCGGGCAT ACGTCGGCCT GCACCCGTTC ACCGCGATCC CACTCGGCAC 1200
GCTGGCCGAC GGGCCGACCC GTGAAGCGCG GAAACGCGCG AAAGGCGCGT TCAACGCAAT 1260
TTGGCAGTCC GGCGCGATGA CGCGCACTGA CGCATACGTC TGGCTGGCTC AGCAGCTCGG 1320
GATCGAGAAC CACGAGGAAT GTCACATCGG CTGGTTCGAT GTCGCGACTT GCGATCGCGT 1380
CGTGACTGTC ATTCAACAGG AGTTCCCCCG ATGACCGATA TGCAAGACCC GCTCTGGCGC 1440

```
GCGCTGAAGC GCCTCGAGCA CGCCGAGCTC AGCGACGACG ACCGCAATCT GCTGCGCCCC 1500
GCATTCGCCG CTCTGCACGG CTCGCAGGCA CTGCGAATCC CCGAGCAGAT CGTGGCGCGC 1560
ATCCGACATC TGGATGCGAC GCTGACTAAA GAGGTGGCTC ATGAAGCGCG ATAGCTTCAC 1620
GCTTTCTCTC GATCTAGGGA GCGAACTGAT CGTCGACAAC TTCGCTGGCG GAGGCGGCGC 1680
GAGCACTGGC CTCGAGCGCG CCTTCGGTCG ACCGGTAGAT GTCGCGATCA ATCACGACCG 1740
TGAAGCACTC GCAATGCACG CGGCCAACCA CCCGCACACC GCGCATTACT GCGAGAGCGT 1800
ATTCGACATC GATCCGGTCG AAATCACGGG AAATCGACCC GTCGGCCTCG TATGGCTATC 1860
GCCGGACTGC AAGCATTTCA GCAAGGCGAA GGGCGGCAAG CCCGTGTCGA AGAAGATCCG 1920
TGGTCTCGCG TGGGTAGCGC TGCGCTGGTG TCTGAAAACG TCGCCCCGCG CGTTCATGCT 1980
CGAGAATGTC GAGGAATTCA TGACATGGGC GGACGTGATC GAGATCAGCC CGGGCAAGTG 2040
GATTCCAGAC CCCGCGAAGA AGGGCGAAAC ATTCGACGCA TTCATCGGCA TGCTGACGAC 2100
GGGCGTCCGC CGCGATCACC CGGCGCTCGC CGAAGCCTGC GAAGTGCTCG GAATTCCGCT 2160
CGACGGACCT GATGCCGATC GCTTGGCAGC AGGCTTGGGA TACAATGTCG AATACCGCGT 2220
TCTGCGAGCG TGCGACTACG GCACGCCGAC GATCCGTAAG CGTCTCTTCG TTGTCGGACG 2280
TCGCGACCAT CTGCCGATCG TCTGGCCGAC GCCGACGCAT GGCGATCCGA AGAGCGCGGC 2340
CGTGCGTGCC GGGAAATTAC TGCCGTGGCG CACTGCCGCC GATTGCATCG ACTGGTCGAT 2400
CTCGTGCCCG TCGATCTTCG AGCGCGATCG GCCGCTGAAG GACGCGACGC TGCGCCGCAT 2460
CGCACGCGGC ATCATGAAGT CCGTCGTGAA TAGCGACGAC CCGTTCATCG TGAAGTTCTC 2520
GCAGAACAGC ACGGGCCAGA CGCTGGACGA GCCCATGCAC ACCGTCATGG CAGGTGCGCC 2580
GCGGTTCGGC GTCGTGGTAC CGCACGTCAC GAAGTTCC                         2618
```

Figure A.8.9 Contig W 862bp

```
AAACTGCGAG GCACCTGCCG CGACGGCGCG CGCGTCGACG AGCCGCTCCA TACCGTCAGT 60
GCTGGCGGCA TGCATCACGC CGAAGTGCGC GCATTCCTGA TCAAGTATTA CGGGAACGAA 120
AAGGACGGCG TGGATCTCCG CGATCCGCTG CACACCGTGC CGACGCATGA CCGGTTCGGC 180
CTCGTCACGA TCCACGGCGA GGACTACGCC ATCGTCGACA TCGGCATGCG CATGCTCACG 240
CCGCGCGAGC TGGCCCGCGC GCAGGGATTC CCGGACAGCT ACGTGCTCGA CCCGGTCGTG 300
AACGGCAAGC CGCTGTCGAA GTCGGCGCAG GTGCGCATGA TCGGCAACAG CGTGTGCCCG 360
GACGTCGCGA CCGCGCTAAT CCGCGCGAAC TTCGCCCACG AACAGCAGCT CGCGTACGCG 420
GCAGTCTGAC CCACAGAGGA CATCACCATG AACRACCAAC AACAGAGCCG CGCTGATGCG 480
CTGACGGTAC TGGCTCGGAT TGTCGAGCTT GACGACGGCA AGAACAGCAT GGCGAACAAA 540
CCGCATCCCG AGCGCTCGCA AGAGTGGAAA GCCTATGCTG CGGAACTTCC AGAATTGCTT 600
ACGAAAGCCC GCGCCATCCT CGCCGCATCC CCTGTCGAGC AGCCCGCAGC AGCGCCGGCC 660
TCTACCAATG AGACAGGCGC GGAAGGGCTC CTGCGGCGCG CACGTGAGGA ACTGTCGTTG 720
GTCGAGTGGG AAGACGRTCC GCCGAACCGC GTTATCGCAC TGTTCGACGA CATCGAAGCA 780
TACGTGTCCC GCTCCCCCGC TATGGCGGCA GCAGCGCCGG CCGACGACGC GCGCGAATGC 840
CTAATGGACG TGGTAAGCCA CC                                         862
```

Figure A.8.10 Contig H 1502bp

```
CATCGGCAAG ACCGCACAGC CGATGCTGCT CTGGGGTGCC GAAGTCAAGT ACACGGTGCC 60
CGAGCTGATC AAGTCGCAAG CGCTCGGCAT GCCCATCGAC TCGCAGAAGG TCGAGGCGAT 120
GAACATGAAG CGCAACATGG ACCTCGACCA GATCGTCTAC TACGGCGATC CGCAGATGAG 180
CTTCACCGGC CTGGTGAATT CGATCGGTGC GGTCGGGAGC GTTTCGAACG TCGCGAACGG 240
CGCGGCCGGC ACGCCGCAGT GGGAAACGAA GACGCCGAAG GAGATCCTCA AAGACGTCAA 300
CGAGATCCTG ACCTCGGCAT GGCAAGCGTC CGGCTGGAAG GTGAAGCCGA ACCGCCTCAT 360
GCTGCCGCCC GCGAAACTCG GCTGGGTTGC GTCGCAGATC GTGAGCGACG CCGGCAACAA 420
GTCGATCCTG ACGTACCTGC TCGAGAACAA CATCTGCACG CAGCAAGGCA CGCCGCTCGA 480
AATCCTCGAG CTGAAGTGGC TGATCGGCGC GGGCGCCGGT GGCACGCAGG GCCAGCTTGG 540
CACCGTGGAT CGGATGGTCG CGTACAACAG CGACAAGAAG TACGTCCAGT TCCCGATGAC 600
GGACCTGCAG CGCACGCCGC TCGAGTACCG CTCGCTGTTC CAGATCACGA CCTACTGGTC 660
GCGTATCGGC CGCGTCGAAT GGCGCTACGG CACGACGGCC GCTTACCGGG ACGGGATCTG 720
ACATGGCGAA GATCAACGTT CTGACGGCGT TCACGATCCG GTTGCTCCAC GAGGGCGAGG 780
AGGTCGTCCG CCGCGTCGAA GCCGGTGTGC AGGAGGTGGA GGACTTCATC GCCGAGCACT 840
GGTACGCGAA GGCGCACACG GGCCCGCTGC CGGAGAAGTC CGGCGATGTG GGTGACTCGC 900
AAGGCGGCGC GACCGATCAG GCTGCGGCGC TCGCCACGGC GAAGTCCGAT CTCCAGGCCG 960
```

```
AGTCGGACCG GCTTGAAAGG CTGCGCACCG AGCTCGATAC GTTCGGCAAG GGGCTGGACG 1020
ACCGCGCGGC CGCGCTCGAC ACGCGCGAAG CTGCGGTCGC GGCGAGCGAG CAGGATCTCG 1080
CCGCGCGGGT CGCGGCCTTC GAGGCAGCTC AGAAGGACGC CGCGGCCGCT GCGAAGAATG 1140
GCGCAGCCGA CGGCGCCAAC CAGAAGTCCA GCAGCGGGAA GAAGGCATAA TGGCCTCCCG 1200
GCGCCGCGCC ATGCAGGCGC GCGCCGGGCA TCCGCATTTT GGCAAGGTGA CACGTGGATA 1260
TCGCCCAGTT CCGACAATCG TTTCCCGAGT TCAACGACAC GACGACGTAC CCCGACTCCC 1320
TCGTCCAGTT TTGGATGACG GTCGCGGTCT CGCTAGTCAA TGCTGATCGG TGGCGCGAGC 1380
TGACGGATCT GGGTGTCGCG CTGGTCACCG CGCACCACCT CGCGCTCGCG CTTAAGGACC 1440
AGAAGACAGC CGCAGTCGGC GGCGTGCCCG GGCAGGTGAC CGGGCCGCAG TCGTCGAAGG 1500
CC                                                              1502
```

Figure A.8.11 Contig G 1568bp

```
CAAGGTGAGC GCGAGCTACG ACACCGCGGC TGTCGCCATC AAGGACGGCG GTTTCTGGAA 60
CGCCACGATG TACGGCGTTC GCTATCTCAG CCTCGCGCAG ATGATGGGCT CGGGCGGCAT 120
TCAGCTGTAA CGCTGCCGCC GCCCATCGGG AGAATCCCAT GGACGGCATG AAAATCGACC 180
GCCTCGACGA GGTGCTGAAG TCGATCAGCA GGCTCGTGCA GAAGGAGGTG CTCGTCGGCG 240
TGCCCGACAG CACCGCCAGC CGGAAGGACG AGGGCGAGCC GCTCAGCAAC GCCGAGATCG 300
GCTACATCAT GGAAAACGGC TCGCCGGCGA ACAACATCCC GGCGCGCCCG CATTTGGTGC 360
CAGGCGTGCA GGACGCGCGG CCGAAGTTCG AGCCGCAGCT CCAGAAGGGC GTCGAAGCGG 420
CACTCGACGG CGATCTCGAT CAGGTCGAAC GCCGGCTGAA GTTGGCCGGT CTTGCCGGGC 480
AAAACGGCGT GCGCGCGAAA ATCAACAGCA ACATCGCTCC CGAACTGGCC GATTCGACGC 540
TGGCCGCGCG CCGGCGCCGC GGCGTCACGC GGGAGAACAC GCTGGTCGAT ACCGGCCAGT 600
ATCGGAACGC GATCACGTAC GTGGTCCGCA AGAAGTAGTA CGATCACTGC GTCGATAGAG 660
GAGTCGGTAA TGACGATAGG CGAAGGAAAC GCGAGCCCTG AGAGACAGTT CGAGCGCATG 720
CGCGCGCGCG TGGCCGAGAT CCATCCTGCG CTGTCCGATG GGCAAATCCG TCACGCCTGC 780
AACGCGATCA GCAACGGCCT CCGTGCATGG AAGGGTGACA TTTATGAGGT GCGCCTGAGC 840
GGTTCGCCTT TTGCCGGCGA GCAAGGATAT TTCCATGACT TTGATGTCAC CTCCGGTGGC 900
AAACGCATAT TCGTGCGTGT AACGGTCGAT TGAGACCAAC GAGACACGAA ACGTAGTCCG 960
ATCATTTCCC AACCCTGAAG GGCCGCCACG TGCGGCCCTT TTTCATTGGA GCTCCGCATG 1020
GCGTTCCTCG ACGTCACCGA CGCCCTGCTC GATCCCGACT TCATGGACAC CGGCCTGCTC 1080
TGCAACCGCA TGACGCAGAC GGTGGACGAC CACGGCCGCG CGCAGAACGC CGTCGCATCC 1140
ATGCCGTTCT CGGCCGTCGT GACGAGCGAC AAGGGCGACA TCCTGCACCG CAACGCGGAC 1200
GGCAGCCGAA TCATCGGTTC GATCACGCTG CACACGATGT TCCGGCTGAT GGACGGCAGC 1260
GCCGGCCACG ACGCCGACGA AGTCGTGTGG GCGGGCCGCA CCTACACGGT GGTGAACGTG 1320
AACGACTACT CGCACTTCGG CCGCGGCTTC GTCTGCGCGA CGTGCGACCT GAAGCCTCTT 1380
TCGGGGTGAC CCCATGAACG ACAGCTCGAC CGGCGGATAC CTGGCGCCAG CCGTCGATGC 1440
GCCGCCGGCC GAGGACGATG CGCTCGACGA TCTCGTCCAT GACCTGATTG CGGGCATCAC 1500
GGCGCTGCCC GGTGATCTCG TGCGGCCGCG CTGGCAGCAG ACCGTCCCGA AGCAACCTGA 1560
GCCGTCCG                                                        1568
```

Figure A.8.12 Contig T 177bp

```
GCACGTTCTA CGGGCCGCGC GCGAAGGGTT ACGCGCAGCG GCTCGCCGAC GGTCTCGCGA 60
TCCCGCAGAA CCGTGAGCAA CTCCAGCTGC AGGACATGGC GTTCGTCGGC GTCGGCGCGA 120
TCCGCGCGGC GCCGGACTTG GTCAACCAGC AGTGGGTGCG CCGGTACGAC ATGACGG    177
```

Figure A.8.13 Contig K 1231bp

```
AAAGGCGGCA GCTTTGCGCT GGTTTTTTCT GGAGGGGAAA CGCGCGGTAT CGATCGCCAA 60
GGCTCCCTCG TTGGCAACGC TGGCAGCGAG CCTGAGCGAA GATCAGCCGT TCACGTTCGC 120
GTTCACCAGT GGAGACAGCA CGCGCGAAGC GAGCATGTAC GGATTCGACG CGTTCTGTCT 180
CGAGAAGTGG CAAGCGGCGC TGGGGACGCG AGGGCTCGAC GTGCTGTCTT TCTCGTTGCC 240
CGTACCTGTC GTCAAGGCCG ATCCTGATCT TCTGCCGAGG TTGTTTGCGG AATTCGCGCG 300
GCGTCTCAAC GCGATTCATG GACACGCCGG CTACGCGGTG AATTTGCCGC CTACGGCGCG 360
CGAAGAGAAC GAGAGCAGCG AATATTTCAT GTCGAATCGA CTCGGACCGG GCCTCGACGT 420
GGGCGATCCT TTTGCAACCG AGGTCCGAAG CCTGATGGAC AACATCAAGA CCGTTGACTG 480
GCTGACGCTG ATCAGCGCTT CGATGGTCGA CCGTGTTGGC GGTGTGAGCG TGCTGAAATC 540
CGAGTTGCCG ATGGATTGGT ATCGGCTGAC GCAGTGCTCG GAGGGCTTGC TCATCCGAGC 600
```

```
CGGCGTTTTG CCTGCCGCGG GCGTGAGTGC GGGGAGTGGC GACAAGCCAG TTGGTCCGCC 660
ACCCGCGTAT GTCGTATTGA ATGCGGCGCT TCGCCACCTT ATCCCCGATA CCGTCTCCAT 720
TCTTCAACGT GGAACCGTGA ACGGCGACGC TCCGGTTTTC AACTCGAAGA CGAGCAGCAA 780
CGCGTGGCTT CGGCGTCTCG ATGTTTCCAG CGATGAACTG TTGGCCGCGA AAGCGGCGGT 840
CCTCGATACG CCGAGACTCT CTGATTCTTC GTCGTAACCC GCGCGAGCCC TGCGTGGCAT 900
GCATTTACTT GACCCACTGC GGGTTACGTT GAATCGCGCT ACATCACGGT CGGCACGCTG 960
GACGGCCGCA TGATCGTGAT GGTTTGGACG CCACGCGGCG AGGCTCGCCG CATTATCAGC 1020
ATGAGGAAAG CAAATGACCG TGAGCAAGCG CGCTATGCAC ACCGATTGGG TTGATCCGGA 1080
CGACGCGCCG GAGCTGACCG ACGAATTTTT TGAGCGGGCG GACGAGTACG TCGGCGACCG 1140
CTTGGTCCGC CGCGGGCCGG GGCGGCCGCT CGGCAGTCAC AAAACCGCGA CGACGATTCG 1200
GCTCGACGAT GACGTGCTCG ACGCCTTCAA A 1231
```

Figure A.8.14 Contig F 2071bp

```
CCGACACGCC GACCAGTTGC AGGCCGTGGC CGACGATGCC GCGCTGGCGC GCGACGAGCG 60
CAATGAAGCG ATCGCCGATG CCGTCACGTT CGACGTGCTT CCGTTCTCGA CCGAGCAGAT 120
CGCCGTGCTC GACGCCGCGC TGCGCCGCGG CTACATCGAG GACGTGTACG AGGTCTGGCA 180
AGGACGTGCT CACCGCCGAG ATCAAGCGGC GCGTCGCCGA TGCTGATCTC GCCGGCGCCG 240
CACCGCGCTT CGACAACGTC GGTTGTTCGG AATGCGGCCG CGGCTTCGGC CCGGGCAACA 300
CCGGGTTCTC CCACTGCGCC GACCACATCG GGTGTCGCGC GCTCGATGAC TGATCAACCG 360
CGCCCGCTAC GGCGGGCAAT CTGACCACAA GGCCGAAACC TGCCCGCCGC AGAGCATCGC 420
GGCGTGGTCG CAGCGTGTGG CGCTGCCTGA CGATGGCCGT CAGTACCGCG AAGAGGATCG 480
ATATGGACAA CACGAAGCAC GGCGGCCCGG CGTTTCCGCT TGCCGATGCA CAGTCGGTGC 540
ATCGCATCGG CGCGGCCGCG ATTGAAGGAA TCACTGATTC GGCCGAGCGC GACCGTGTCT 600
ACATCGAAGC GACGTCTCGC GCTTGCGCCG GCATGACGCT CCGCGACTAC TTCGCGGCGA 660
AGACGATGCG GCCACTGACG CTCTCGATGA AGAGCGCCCG AGAGAAAGAA ATGCACAGCA 720
TGGCGCGAGA GGCATACGCG ATCGCCGACG CCATGCTCCG CGCGCGAGGA GCCTGACATG 780
AACAAGAACA AACACTCAGT CGGACAAGCA TGGCTGCTGC GTGCGAAAGA ACTGGCGCAG 840
GAGCGAGCTG ATTCGAGCTT CGCGTACGGC GAGCTTCCGT CGGATGCCGG TGACGGTGCA 900
TCGGCCGATG CTTACGCCCG GTCGCTCGCG GCCGATCGAT CGCTCGACGC GCACCTTCAA 960
CCGATGGCGC ACCTGATCGA CGCGCTGCAC CTGGTCGCGT CGAAGACCGT GCTCACGTCC 1020
GGCATCCGCG CCGTCGTCGA CGATGCACTG GTCGAAGGTC GAGGTTTTCC GCGACGCCCG 1080
TGGCCAGTGT CCGATCCGGT TCGCCACATC ACCATCGACC GGAGTCGACC GATGAGCCCC 1140
CTCAAATCCC CCATCCAATG CGGCGACCTT GCTGAAAAGC TGATCGCCGA CTACGTGCGC 1200
GAATCCGGTG CGTATGGCAA CCCGAACGCG CTCGCGAACG TGATCGAGAT GCTGATCAGC 1260
AAAGCGGCGC TCGGCATCGC CATGGTCGGC AGCGAGGCGA TCGCCCAACA GATCCTCGAT 1320
CGCACGAAGG CACAACGTTG CGACGTTCGC AGACCGGAAC CTTCGGAGGG GGCACTGATG 1380
CGCACGCCGT CCAACAGCCT ACAGCCTGTA TTTTTCAAAA ACAAACGATC GCTTGTATTG 1440
CGCATACGGT ATGTGCTCGA AGGATCGGCA TGGGCGGCCG CCTACGGCGT CGCGATCGGG 1500
TTCCTGTGGT ACGGCGCGCT CGTCGCCGGA CCGTACCTGC GGAGCCTCGG ATGAGGGCCC 1560
TACTTCGCAA AATCGCCGAG CTGTTCGCCC TCTGGATCGT CATTACGGCA GGACTGTTCC 1620
TGTTCGTATG GCTCGTCGTG CCGACGCTGA TCAGTGCGCC GGAAGACGCG CCCGTCACCG 1680
TTTCGACGAG GTCCACATGA GCCGCTTCAC CGATCACGCC CAACGCTTCG AAATTCATCA 1740
CCCGCGCGCT GCCCGGGTGC TGGTGCTCGC GATATTCGTG GCCGTCGCAC TCCTCGCCAT 1800
AGCCATCGAC AGCATCACGA AGCACTTCGG AATCTTCTAA GAGGCACCGC CTTGCGCTCC 1860
CCGGAGCGTA TTCCCTGCAC TTCCCGATCT CGCTCGCGCT CATCGGTCGC TGTCGGCGCG 1920
CGTGTCCACC ATTCATCGGA GCACCACCAG ATGAGACCAT GACACCTGCC CGTCGAGTCG 1980
TCGCAGATCC CAGGAAACAG GCTATGACGA TGAAAACCGA AACGCTTCGC AGGGTTTGGT 2040
TCATGGAATC TGCGTTGCGA GAGTACTGGA T 2071
```

Figure A.8.15 Contig J 1366bp

```
TCGACATGCT GACGCTCGAT CAGCAGCACG AATCCGACTA CGGCGACGCG AAGGACTGGG 60
CGAAGCAGAC GAGCGGCCGC CGCGAGGATC GCACGATCCG CGCGCTCTTC TCCGAGGAAC 120
TCGGCGCGGT CGTCCAGGTG CGCGCGGCGG ACCGCGACGC GGTGCTCGGC GCGCTGCGCG 180
AGCATGGCCT GTCCGCGTGC TCGCACGTGA TCGGCGCGGT CAACGAGACC GACGCGATCG 240
AGGTGTACCG CGACGCGAAG AAGATCTACG AGGCGCCGCG CGTCGAGCTG CAGCGCGCAT 300
GGAGCGAGGT GAGCTGGCGG ATCGCGCGGC TGCGTGACAA CCCGGCGTGC GCGGACGCCG 360
AGTACGACGC GATCCTCGAC GCCGGCGATC CGGGCCTCTC GCCCGTGCTG ACGTTCGATC 420
```

CGGCGGAGGA TGTCGCCGCG CCGTTCATCG CGACGGGCGC GCGTCCGCGT GTCGCGATCC 480
TGCGCGAGCA GGGCGTGAAC TCGCATCTCG AAACCGCCTA CGCGTTCGAC CGCGCGGGCT 540
TCGACGCGCA TGACGTCCAC ATGAGCGACC TGCTCGCCGG CCGCGCGACG CTCGCCGATT 600
TCGCCGGCGC GGTCGCGTGC GGCGGCTTCT CGTACGGCGA CGTGCTCGGC GCGGGCGAGG 660
GCTGGGCGAA GACGATCCGC TTCAACGACA AGCTCGCCGA CATGTTCGCC GCGTTCTTCG 720
CGCGCCCCGA CACGTTCGCG CTCGGCATCT GCAACGGCTG CCAGATGATG TCGAGCCTCG 780
CGTCGATGAT CCCGGGCGCG CAGGCGTGGC CGAAGTTCAC GCGCAACAAG TCCGAGCAGT 840
TCGAGGCGCG TTTCTCGTTC GTCGAAGTGC AAAGCTCGCC GTCGATCTTC TTCGCGGGCA 900
TGGAAGGCTC GCGGATTCCG GTGGCGGTCG CGCATGGCGA AGGCTACGCG GACTTCTCGC 960
AGCAGGGCGA TCAGAGCCGC GTCGCGGTCG CGATGCGCTA CGTCGATCAC CGCGGCGACG 1020
CGACCGAGCG CTATCCGTTC AACCCGAACG GCTCGCCCGC GGGCATCACG TCGGTGACGA 1080
CGGCGGACGG CCGCTTCACG GTGCTGATGC CGCACATGGA GCGCGTGCAT CGCACGGTGA 1140
CGATGAGCTG GCATCCGCAG GGCTGGGGCG AAGCGAGCCC GTGGCTGCGC GTGTTCCGCA 1200
ACGCGCGCCG CTGGATCGGC TGACGCGATG ACCGCCGCGG CGACCGCGCC GGGCGGTGCC 1260
GTCACGCTGC GCGGCGAACG CGCGGGCGAC GCCGCGGCGC TCGCGCGCGT GATCGTCGCC 1320
GCATTCGCGG ATGAGCCGCA AGGCGGGCAA TTCGAGCGGC GCATCG 1366

Figure A.8.16 Contig Px 434bp

CGGACGACGA GATAAAGCTC TTGAGAGAAC TTCGCGAGAT TCGCAATCGC GCTGCACATT 60
CTGTTAAGGG CCGACCAACG CCAGATGAGG CCGAGCGATT TGTTTCGATT GTTCGCGCAC 120
TTGAGGCTGC TTGGATTGCT CGATTGGCGA GCGCGGAGCC GAGGTAACTT GTCAAATCGA 180
AAGAGATGGA GAGGATCTGC GAGTCCAGCC TCTGCGAGCG CAATTACAGG GCGCCAGCCC 240
GGTACGACGC ACAGAGAAAC CATACGCCTC CGATTGTGGT GCCATAGGCG AGGATTATCC 300
CGAAAGCTCG AGCGAGACGC CCGTTCGGCC GAGCGCACGC GGCGAGTAGA TCATTGTCGA 360
AAGCGACTCT GTTCATGATG GTCTCGTGTC GTCGCAATTA CCGACGGTAT GCATGGGTCG 420
AGGGCGGTCA GGTC 434

Figure A.8.17 Contig N 493bp

GCTGTAGTGG TTCAATGCCC TTGTGACGGT TCGAAAAGAA GGTAACGTGT CGTCTGGAGG 60
GCCAAACCAT GGTCAGGCGT GCGTTTTCCG AGGAATTCAA GGAAGAAGCG ATACGACTCG 120
TCGTAGAGCA GGGTTACCCC TTTTCCAAGG CGTGTGGGGC GGTTGGAATT GGGGAGACAG 180
CCCTGCGTCG CTGGGTCGCG CAGTGGCGGG CGGCGCATAG CTTGGATGTG CCCAGTCCTG 240
CACAACTCAG TGCAGACGCG CGCCGGATCA AGGAGCTAGA GGCCCGCGTA GCCGAACTCG 300
AACGGGAGCG TGAAGTTTTA AAAAAATCCA CAGCCTTCTT CGTCAAGGAG CTCGAACGCT 360
CCTGGAAGTG ATCCAGGCGA TGGAGAAGGC ATACCCGGTA GCGCTGATGT GTCGACTGGC 420
TGGTGTAGCA CGCAGCAGCT ATTACGCATG GAGGTCGCGC CAAGGCAGAG CCAATCGCGA 480
CGCGACGGTG CTG 493

Figure A.8.18 Contig 7 1772bp

GCGCGCCCGC CGCCTTCGCG ATCGCGGCGA TCTCCTTCAG ATGGCGGTTC CACACGGTGT 60
TCGACGTCTC GCCATGAACG ATCGTGACGA TCTCGGGACG CTCGCGCGCG ATCGCGTCGG 120
CGATTTCGTC GAGGCTCGCG ACCGAACGGT CGGGCACGTC GATCGTCGCG ACGTCCGCGC 180
CGACGCGCAT CGCCATCTCG GCCATCCGCG CGCTGAAGAA ACCGTTCTTG ATCGACAGCA 240
CCTTCGTGCC GGCCCACGCG AGATTCGAGA CGGCCATCTC CATCGCGGCG GAGCCCGGCC 300
CCGCGACGCC CAGCACCCAC TTCGTGCGCG TCTGGAACAC GTAGCGGGCC ATTTCCTTCA 360
CCTGCTCGAT GATCTTCGCC ATCGTCGCGC CCAGGTGGTT GATCACGATC GTGTTCGCCT 420
TCGCGACCGC GGCGGGAATC GGCACGGGGC CGGCGCCCAT CATGAGCAGC GGTTCTTCGG 480
GCAGGATCGC GTCGAGCGGG ACGACGACGG GACAGGGAAT GGGCGAGTGA TCGGTCGTGG 540
ACATGGGCTA ACGAAGCACG GCGGGTTGGA AAAAGGGGGG AATCGGTCGA TCATTCCGCG 600
AACCGTTGCA ATGCGCAAGT GTTTTGATCG ACATACGCGT CGGGTGCTTG CGCGGCGTCG 660
GGCGGCGCGC TCGCACCGGC GGCGGCGCCG ACCGGCTCCC CCTGCCGACG ACGAAGGGCG 720
CACGCCGCGG CCGGCGCGCC CGTGGCGGAG GCGCGTTCCG GCCCGCCGCC CTACGATATT 780
TCCGAGGTTG ACGCCCCCGC CCGTCGCGAG AAAGCATCCG CGCGCCGCAG ATTTTCGCGC 840
CGTGCGCATC GGAGCATTCC GGCTCGCCCG GCCCGAATAG GCCAAAACGG CATCAAATCG 900

ACTCCCTATT AAATCGAAAA ATCATTTTTT GCAAGCCGGG CCGTCAAACT TTCCGTAAAA 960
GCTTGCATTA ACCATTCGAT GCGGTGCAGC AAACTTATTT GTTCCGCACC GGAAAAGAAC 1020
CGGACACCTC ATTGCGACTC GCACACAAAA ATCCCCAAAA AAATACGCCA ACGGCATCCC 1080
CGCGCGAGAC ATGCCGCACT CCGAAGCAGC GGAGTCCCGG AATGGGAACG GTCCCGCACC 1140
GGCGCGCCGG CCGGCCGCCT TGCCGGGATT AAATCCGGCG CAATCGCCGA ATATCGCGCA 1200
CGCGAATCGC GGAATCGGTT CACGCCCCAT AAATAGGGAT CGCCGCGGGC CGGCGGCCGA 1260
TTTATTACAT TCGGGATGAA ATAATTATTT AGTTGACTGA TGCTGTCAGC TTATTGAACA 1320
ATTTGCCCAG CCTTCGATGT TTCGTTATTT AAGACAAGCC CAGTGCTGAT CCGACAGCCT 1380
GCGGCTCGCC AGATCGAAGG CGAAAAACAC AACATAATCA TTTAATGGGC CAACGAAAGC 1440
TATGAACCAA ATCCAGACCA TGCGTGTCTT CGTATGTGTT GCCGAACTGC AGAGCTTCCG 1500
CCGGGCGGCG CGGAAACTCG GCGTATCGAA TGCGCTCGTC ACGCGTTCGA TCGCGATGCT 1560
CGAAACACAT CTGAACACGC GTCTCATCCA TCGTACGACA CGCAATCTGT CGCTGACGGA 1620
AGCGGGCATC CGCTATCTCG ACGGCTGCAG GGCGCTCCTC GAGGAGTTCG ATCAGCTCGA 1680
GGCGTCCGTC GAGCGGGCGG TGCACGAGCC CGTCGGCACA TTGCGCGTGG CCGTGTCCGG 1740
GTTGCTGTCG CCCGGCCGGA TCACCCCGCT CG 1772

Figure A.8.19 Contig P1 1201bp

CGACCCACGG CAGCGCGCGA ATCGCGAGGC CGCCGAGCCA GTAGCCGAGC GGCCCGTCGG 60
ACGTGACGAA TTTGCCGACG AGATTGGGCA GCAGCCAGTC GTGCAGGCCG CCTTGCGCCA 120
TCGTCCACAT CACGCCGAAG CCTGCCGCGT CCTCGTTCTT CCACGGATCG CGGCCGAACA 180
GCCCGAACGC CGCGTAGACG ATGCAGAGCG TGAGCAGCAG CCAGCGCGGC AGCGCGCGGG 240
TGGCGGAAGC GGTAAGGCGA ACGACAGGCT TCATGCAGAT GGGGGCAAGA GGATGGACGA 300
CATGCCGGCG CGCGAGCGGC GGGCGGTTTC GGGCATCCGG CATTGTAGAC GCGCCGGCCG 360
GCACGCGTCA GCGAATCGGC GGGCCCGCGC GGCACCGGTG GCGCGCGGGA TCGACAGGGA 420
TAAAAAAAGG CAGCCTGGGC TGCCTTTCTT GCGGAACCGG CCCGGCGCGA AGCCGGATGC 480
CGATGCGGTC GCGAAGCTTA CTTCGCGGCC TTGCCGCTGG CGCGTGCGCC GAACTTGTTC 540
TTGAACTTCT CGACGCGCCC CGCCGTATCC ATGATCTTTT GCTGGCCGGT GTAGAACGAG 600
TGCGATTCCG ACGACACTTC GATCTTCGCG AGCGGGTAGG TCTTGCCTTC GAATTCGATG 660
GTTTCGCGCG TCTGGATCGT CGAGCGCGTG ATGAACTTGA AGCCGTTCGA CATGTCTTGG 720
AAGACGACTT CGCGGTAATC CGGGTGAATG CCTTGTTTCA TGGTTTTTCC TGTGTAAGAG 780
CGGTAGCCAG CCCGCCGCTT GGCGCATCGC GGCGCGCGTC GCACGATGTC CGGACACGGC 840
CGGCGCATTC GTACTGGGCG CACCAGAATC GACGCCCGTC GGCGCTGCGC GGTGCGAAAG 900
CCCGGCGCGA GCCACTTGCC TAAGGTCGAA AGCTCGCGAT TATGCCAGAA AATCAGGCGC 960
TTGGCGACTT TTTCGTCAGC CGTGCGCGGC GGAGCCCGCG AGGCACCCCG TGCGCGCGGG 1020
ATCCTGCCGG TAATAGCGGG CGAGCAGCCG GTAAAGCTCC GGAAATTCGG CTTCGAACGG 1080
GCGCGGCTTG ACGAAGAGCG CCTCGCTGCA GACCGCGAAG AATTCCGACG GGTGGTCGGC 1140
CGCGTACGGG TCGATCAGCG AATCGCGCTC GAAGCGCGCC CACGCGCGCT CCGGCACCGC 1200
G 1201

Figure A.8.20 Contig I 1390bp

GACATGGAGG CCCAGAATAC CCTCCTTGAC AGTCTTGACG TGCGCAGCTC AGGGGCATGA 60
TGTGACTGTC GCCCGTACAT TTAGCCCATA CATCCCCATG TATAATCATT TGCATCCATA 120
CATTTTGATG GCCGCGCGGC GCGAAGCAAA AATTACGGCT CCTCGCTGCA GACCTGCGAG 180
CAGGGAAACG CTCCCCTCAC AGACGCGTTG AATTGTCCCC ACGCCGCGCC CCTGTAGAGA 240
AATATAAAAG GTTAGGATTT GCCACTGAGG TTCTTCTTTC ATATACTTCC TTTTAAAATC 300
TTGCTAGGAT ACAGTTCTCA CATCACATCC GAACATAAAC AACCATGGGT AAGGAAAAGA 360
CTCACGTTTC GAGGCCGCGA TTAAATTCCA ACATGGATGC TGATTTATAT GGGTATAAAT 420
GGGCTCGCGA TAATGTCGGG CAATCAGGTG CGACAATCTA TCGATTGTAT GGGAAGCCCG 480
ATGCGCCAGA GTTGTTTCTG AAACATGGCA AAGGTAGCGT TGCCAATGAT GTTACAGATG 540
AGATGGTCAG ACTAAACTGG CTGACGGAAT TTATGCCTCT TCCGACCATC AAGCATTTTA 600
TCCGTACTCC TGATGATGCA TGGTTACTCA CCACTGCGAT CCCCGGCAAA ACAGCATTCC 660
AGGTATTAGA AGAATATCCT GATTCAGGTG AAAATATTGT TGATGCGCTG GCAGTGTTCC 720
TGCGCCGGTT GCATTCGATT CCTGTTTGTA ATTGTCCTTT TAACAGCGAT CGCGTATTTC 780
GTCTCGCTCA GGCGCAATCA CGAATGAATA ACGGTTTGGT TGATGCGAGT GATTTTGATG 840
ACGAGCGTAA TGGCTGGCCT GTTGAACAAG TCTGGAAAGA AATGCATAAG CTTTTGCCAT 900
TCTCACCGGA TTCAGTCGTC ACTCATGGTG ATTTCTCACT TGATAACCTT ATTTTTGACG 960
AGGGGAAATT AATAGGTTGT ATTGATGTTG GACGAGTCGG AATCGCAGAC CGATACCAGG 1020

ATCTTGCCAT CCTATGGAAC TGCCTCGGTG AGTTTTCTCC TTCATTACAG AAACGGCTTT 1080
TTCAAAAATA TGGTATTGAT AATCCTGATA TGAATAAATT GCAGTTTCAT TTGATGCTCG 1140
ATGAGTTTTT CTAATCAGTA CTGACAATAA AAAGATTCTT GTTTTCAAGA ACTTGTCATT 1200
TGTATAGTTT TTTTATATTG TAGTTGTTCT ATTTTAATCA AATGTTAGCG TGATTTATAT 1260
TTTTTTTCGC CTCGACATCA TCTGCCCAGA TGCGAAGTTA AGTGCGCAGA AAGTAATATC 1320
ATGCGTCAAT CGTATGTGAA TGCTGGTCGC TATACTGCTG TCGATTCGAT ACTAACGCCG 1380
CCATCCAAGA 1390

Figure A.8.21 Contig 11RF-smF 1053bp

GGCAGAAGTT CGTGATCGGT GACGCTGACT TGCGAGTCCT CCGAGAAGCC CAGCGTTCGC 60
TGGTCCCGCT GAAGACAGCG GCTCGATATG CCGGGATGTC CACGAGACGT ATCCAGGCAC 120
TGGCAACGGC GGGCCTGCTT GCATCAACCG ACACCCGGAT CGACACGAGG TCCGTTGATC 180
GTCTGCTCAA CGACATCGTG GCAGCCAGTG TCCGAAACCT ACGGGCGTTT GAGGATCCAG 240
TCAGCGTAGC CGATGCCTTG CGTCTGTACA TTCCTGTTAA AGCCTCAGCG GTATTCTTTA 300
ATCGACTCGT GAAAGGTTCC GTGCGCCTTG TCTTCGAGCC GGACAATGCA CCGATACTTC 360
GGAACATGTT CGCCAACCGC AACGAGGTGC GCTCCGCTGC AGAGATGCCG GTCGAACCGA 420
GCTCCCAGAT ATCTATTGTT GAGGCCGCCC GTCGGCTCGG CGTCAAACAG GAGGTGATGT 480
ACCACCTGAT CAACAAAGGC CTGGTCAGGA CCCGAATGGG CAAACTGGGG CACCGGGCAG 540
CGAGAGTCGT TGATGTCGAT GACCTACGGA CGTTCGCCGA ACAATTTCTG CCGCTGTTCA 600
CCGTAGCGAA GGCGACGGGA ATTTCCGCTC GCGAAGCGCC AGATTGGGCC AGGCTGCATG 660
GCATTGAGAT CGTCACTGGA CCATCAGTCG ACGGCGGACG GCAGTACTGG ATCCGCAGGC 720
AGGCGTGGAA ATTTCCTGCA GCGTAGCGTG TCAAGCGTGA AGTATTGATT TTCGTGGGTC 780
GTGTGCCCCC TGTCCTGCAT CAGGATGCGT ATCGACACTG TCGAACGACC TGGATTGCCG 840
CCAGCTAGTC AGGTCAGCAA CCAGACCTTT GATTGTGAAA CCATGAAACT GCTCATTCTG 900
TCTGATCTGC ATAATGAATT CGAGTCGATC GCAGCTGATG TGGAGGCTGT CGCTCGGGCC 960
GACGTCGTCG TGCTGGCCGG CGATATCCAC ACGAAGAATC GGAGCATTGA ATGGGCTCGG 1020
GCATTTGTCG GCGATCCGGC GAAGCCCGTC ATC 1053

Figure A.8.22 Contig S 219bp

GACCTGACCG CCCTCGACCC AATCATCGAC GGATTTCACC TAACACGAGT CTTAATGGAC 60
GGCGGTAGTA GCCTCAACCT GCTCTATCAG GATATAGTGC GGAAGATGGG CATTAATCCC 120
TCACGAATCA AACCCACAAA GACTACCTTT AAAGGAGTCA TACCAGGCGT AGAGGCCCGT 180
TGTACGGGCT CAATCACGCT GGAGGTGGTC TTCGGTTCT 219

Figure A.8.23 Contig L 816bp

CGACCAGGAC GGCCCACAGG AGGATGGTGG TCACGGCCAA GGCGGTTTCG ACTTCGACGT 60
CGCCGGCCTC GTGCGCGGTA TCCGCGAGGA CATCGAGTCC GCCAAGACGC CCGAAGACCT 120
AGACCTGGCC CGTAGCGCGA TCAGCGGTGT GCCGGATGAA ACCGCCAAGG CCGAACTGAA 180
CGCCCTCGCC TCGGCGCGTA TGCGCGCTAT CACCGCCGCG GCGGAGCAAT CGGCCGCCGG 240
CAAAGCGACC GCCCAAACGA CCGCGCCGGC CGGCCGCCGG ACGCGCGGCC CGATCAGCGC 300
CGACTAACCC ATCCAATTCT CCGCCAAGGA TTTCGACATG ACCGACAAGA ACGTCCTCCA 360
TATGACTGCC GAGACGATCG GCAAAGACCT GCTGTCCGCG CTCGTTACCG AGATCAAGCT 420
GATGCCGGAC CTCTGGGTCA AGCTGTCCCA GAACAAGCAG AACGACGTCA TCGACCGGCT 480
GCGTGCGCGC GTCGAGCACA ACGTCAAGAT GGCGACGCAC CTGATCGCGA GCGATGGCCG 540
TGTCGTTGTC CAAGGCGACC TCGTGCAGAT CACGATCAAG GACGGCGTCA AAGCGGCGGT 600
GGAATTCAGC AGCGCCGCGC CGAATCTCCA TGACCTGTAT GACGCGCAAG GCAAAGCCGT 660
CCTGATCGTC GTCGCGAACG CTGCCGCGCA CACCGGCGGC ATGGACGAGA TCCGCGGCGA 720
GTCAGATCAG CGCGGTCTCG ATCTCGGTCG CGAGTACACG GACGACAACG GCGACGGCAT 780
GGACGGCCAG CGGCCCGACG GGGACTACGT CGTCTA 816

Figure A.8.24 Contig DX 968bp

CGACGCCGAG TTCCGCGAGG TGCCGAAGCT CGGCGACGGC CCAACGCAAG CGCAGATCGA 60
CGAGCAGCAT CAGGCCGGCC GGCAGGCGGC CGCCGACGGC AAGCCCGAAA GCGAATGCCC 120
CGTGATGGCC GGCGAGCTGT GCATCGCATG GGTGAAGGGC TGGAAGGAGT GGCATGAGGA 180
ACAAGCCGCC TCCGGTAACG AGGATCCGCT GTACGCCCAA GTCGAAGCGT TCGTGATCGA 240
GCAGCAGAAA GTGTCGATTT CGAGCGTGCA GCGCCAGTTC AAGATCGGCT ACAACCGCGC 300

```
CGCGCGGCTG GTGGAGCTGC TCGAAGCCAA GGGCATCGTC AGTGCGATGG ATTCGGACGG 360
CGGCCGCACG GTGCTGCGGC CGCGCGGACC GCAGGGAGAG GAATCGTGAA AATCACCGAC 420
ATCTACGTGG CGAACGTTCT CGGGATCCGC ACGGCGGACC TCCGGCTCGC CAAACCCGTC 480
TCCCTCTTCA CCGGCCCGAA CGGCGCTGGC AAGAGCAGCC TACAGGAAGC CGTGCGCATG 540
GCGCTCACCG GTGACACGGT GCGTGTCGCG CTGAAGAAGG AATACGGCTC GCTCGTCACC 600
GAGGGAGCCG ACGGCGGCCA GATCGTGGTT GCGTGCGGCG AGCAGGCGAA CAGCGTCATG 660
CTGCCGTCCG GAAAGCTGAA GCGCGAACTC GCCGAGGACC CGCGCCTTCC GCTGGTACTC 720
GACGCGCAGC GATTCGCGCA CCTCGGCCCG GCTGAGCGGC GCGCATTCCT GTACGACCTG 780
ATGGGCGTGA AGATCGGCGT CGACGAAATG CGCGCCCGGC TCCTGGACAA GCTCGGGTTC 840
CGCGCCGATG CGGTGCCAGC TCCGGCCGCC GCGCGGCTCG CTGCCATCAC GCCGATGTTG 900
CGCGCCGGCT TTGAAGCGGC GCAGAAGGAA GCCGCCGACC GCGCACGCGG CGCGAAGCAG 960
TCGTGGCG                                                        968
```

Figure A.8.25 Contig 2FR 1689bp

```
GCCGAGATGG ATCATCTTCG TGCCGGTGTC GGCCTGCTGG CAGTGGTTCG ATACGGCGAT 60
CGAATGGAAC TCGCCGCTCG ATTCGTCGCC GCGCAGCACG ACGCTCGGAT ATTTCCACGT 120
GATCGCGGAA CCCGTCTCGA TCTGCGTCCA CGCGATCCGC GAGCGCGCCC CCGCGCACAG 180
GCCGCGCTTG GTCACGAAGT TGTAGATGCC GCCCACGCCG TCCGCGTCGC CCGGATACCA 240
GTTCTGCACG GTCGAGTACT TGATGTGCGC GTCGTCGTGC GCGACGAGCT CGACGACGGC 300
CGCGTGCAGC TGATGCTCGT CGCGCTGCGG CGCGGTGCAG CCTTCGAGAT AGCTGACGTG 360
GCTGCCGGGC TCCGCGATGA TCAGCGTGCG CTCGAACTGG CCGGTGTTTT GCGCGTTGAT 420
CCGGAAGTAC GACGAGAGCT CCATCGGGCA GCGCACGCCT TGCGGCACGT AGATGAACGA 480
GCCGTCGGAG AACACGGCCG AGTTCAGCGC CGCATAGAAA TTGTCGGCGG GCGGCACGAC 540
GGTGCCGAGG TAGCGCTCGA TCAGTTCGGG GTGATGCTCG ACCGCGTGCG AGAACGAGCA 600
GAAGATCACG CCCGCTGCGG CGAGCTGCTC GCGAAACGTC GTGCCGACCG ATACCGAATC 660
GAACACGGCG TCCACCGCGA CGCCCGCGAG CCGGGCGCGC TCGTGCAGCG GCACGTTCAG 720
CTTTTCGTAG TTCGAGCAGC TTCGGTCGAC TTCGTCGAGG CTCTTTGGCT GATCCTTCAG 780
CGATTTCGGC GCCGAGTAGT ACGACTGCGC CTGGAAATCG ATCGGCGCGA TGCGCAGCTT 840
GCCCCAATCG GGCGGCGACA TCGCGAGCCA GCGCTCGAAT GCGGCGAGCC GCCATTTCAG 900
CAGGAACGCC GGTTCGCGCT TGCGGCGCGA GATCTCCCGA ACGACGCCCT CGTCGAGGCC 960
CGGCGGTAGC GAATCGGATT CGATATCGGT CACGAAGCCG TGCCGGTACG TTTGTTCGAG 1020
CGCTCTGGCG AGCGGCGTGG GTTGATCGAC GTTGAGCATG GGCGAATCCA TGTAGAGCAG 1080
GGCGGTAGCC TCCGCGCGGT TCGACGTGCG ACGCGCGGCG CGACTCGGTT CGGTCGTTGG 1140
CCTTAATTGC TGTATATGGT ATGCATCTAT TTGGCGGGCC GCAAGCGTGC GTTCGGAGAC 1200
CCTGCTTCGA GTCGTGTCTT TGCCGCAGGG GCGCGGCGGG GCGCGGCCGG CAAACGAGGC 1260
GGCGCCGCCC GCGCGTGCGA CAAGAGGCCG CAGTCGCGCG ACGGGCAGCG GGGGTCGCAT 1320
CAGCCACGAG GATGCGTGTT CGAACGACGC GGCGGCCCGC TGGAGTTCAC GGTGGGGATG 1380
ACGCGCGAGA CTGCTTCGGC ATCCTTCTCG ATCGGGGACG GATGCGCGAG GGCGACGATG 1440
CGCCGCGGTG CCGGCATGCT TCGCACCGCG GTCGCGGCAG TAGCGGCTGC CGCCGCATCG 1500
CACGTCGCCC GCGGCGCGGC GGCGCGTTCG AACGCCGCGC GGGCAGGTCA ACCGGCCTTG 1560
CCGTCGATAC GCGGCGCGGT CAGATAGCCG GTGTAACGCA ATGCATACAC CGCGAACGCG 1620
GCCACCCAGC ACGCGCCCGC GACGTCGATC CAGACGCGCG TCGCATCGGG CGCGATCCAC 1680
GGTGCGAAC                                                       1689
```

Figure A.8.26 Contig Dr-15b3F 1279bp

```
GATCGGCAGC GGCAAGGACA TCAATGTTCA AGGTAGTACG GTCGTCGGTA CGCACGATGT 60
GGCGCTGAAT GCGGCGCACG ACGTGAACAT AACGACGTCG CAGGACACCA GCCAATCGTC 120
GACCACCTAT CAGGAACAGC ACTCGGGTTT GATGTCGGGC GGCGGTCTGT CTTTCTCGGC 180
CGGCAACAGC AAGCTCGCGC AGCAGAATCA ATCCTCGAGC GTCACGAACA ACGCCAGCAC 240
GGTCGGCTCG GTCGACGGCA ATCTGACCGT CAATGCGGGT AATACGCTGC ACGTGAAAGG 300
CAGCGATCTG GTCGCGGGCA AGGATGTGAC CGGGACGGCG GCGAACATCG TCGTCGATTC 360
GGCCACCGAC ACCACGCATC AAGCGCAACA GCAGCAGACG AGCAAGAGCG GGCTGACGGT 420
CGGCCTGTCG GGCTCGGTGG GCGATGCGAT CAACAATGCG ATCAGCGAGA CGCAGGCCGC 480
GCGCGAATCG GCGAAGGATA GCAACGGCCG CGCATCGGCA CTGCATAGCA TCGCGGCGGC 540
CGGTGATGTG GCATTCGGCG GTTTGGGGGC CAAGGCCCTG CTGGACGGCG CGAAGGGGCC 600
GCAGGCCCCG AGCATCGGCG TGCAGGTCAG TGTCGGTTCG AGCCATAGTT CGATGCAGTC 660
GTCGGAGGAT CAGACGATTC AGCGCGGCTC CAGCATTAAC GCGGGCGGCA ATGCGAAGCT 720
```

GATCGCGACG GGCAACGGTA CGCCGAAGGA CGGCAACATC ACGATCGCGG GCTCGAACGT 780
GAACGCGGCC AACGTGGCGT TGGTCGCGAA CAACCAGGTC AATCTCGTCA ACACGACCGA 840
TACGGACAAG ACGCAGAGTT CGAACTCGTC GTCGGGGTCG AGCGTCGGCG TGTCGATCGG 900
CACGAACGGC ATCGGCGTCT CCGCATCGAT GCAGCGCGCG CACGGCGACG GGAATTCGGA 960
CGCGGCGATC CAGAACAACA CGCACATCAA CGCCAGCCAG ACCGCGACCA TTGTCAGCGG 1020
CGGCGACACG AACGTGATCG GCGCGAACGT GAACGCGAAC AAGGTTGTGG CCGACGTAGG 1080
CGGCAACCTG AACGTGGCGA GCGTCCAGGA CACGACCGTA AGCGCCGCGC ATCAGTCGAG 1140
CGCGGGTGGC GGCTTCACGA TCAGCCAGAC CGGCGGCGGA GCGAGCTTCA GTGCGCAGAA 1200
CGGCCACGCG GACGGCAACT ATGCGGGCGT GAAAGAGCAG GCCGGCATCC AGGCGGGCTC 1260
CGGTGGTTTC GACGTGACC 1279