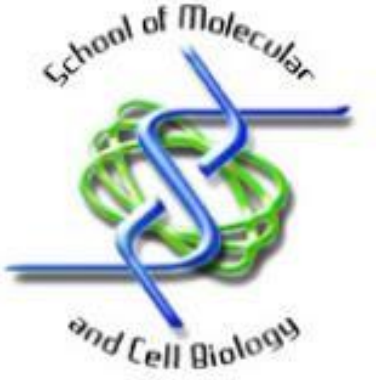




# Isolation and Characterization of Novel Antimicrobial Genes against Mycobacteria through the Exploitation of Mycobacteriophages Genes

**Tsz Hoi Felix Tam**

**0311424k**

**Dissertation submitted to the School of Molecular and Cell Biology, Faculty of Science, University of the Witwatersrand for the Degree of Masters of Science.**
**March 2012**

# Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Masters of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.


09/03/12
Date

# Abstract

Mycobacterial infections are responsible for some of the most well known disease, including Tuberculosis. Reported cases of infections caused by mycobacteria that are becoming increasingly resistant to traditional antibiotics are on the increase. This calls for a new approach in developing new drugs that can act on novel antimicrobial targets. One such alternative involves the use of bacteriophages and the study of how they interact with their hosts. Their diversity also suggests that there are many different phage-host interactions acting on multiple targets that are currently still unknown. Eight phages were isolated and characterized. Genomic libraries were constructed on four of these phages and screened for antimicrobial activities using *Rhodococcus erythropolis*. Six clones were further analyzed, and 15 ORFs were predicted with 8 ORFs being assigned functions. These genes with similarity to proteins in the database suggest that they are involved in membrane integrity and DNA metabolism. These clones were further tested on *Saccharomyces cerevisiae* to determine whether they have any effects on eukaryotes. The lack of inhibition in *S. cerevisiae* suggests these phage products are confined to act only in bacteria after millions of years of co-evolution with their host counterparts, and further studies into these genes will continue to shed light on bacterial genomics.

# Acknowledgements

I would like to express my appreciation to those who have helped me throughout my studies

To Prof Eric Dabbs, for his patience and guidance throughout his supervision.

To Dr Youtaro Shibayama, for his patience in demonstrating to me all the laboratory techniques required, and for the vectors and reagents he has kindly provided from time to time.

To the National Research Foundation, for funding.

To my parents, for their continual support.

# Table of Contents

# List of Figures

# List of Tables

# List of Symbols and Abbreviations

| | |
|---|---|
| A | Adenine |
| ADP | Adenosine diphosphate |
| ATP | Adenosine triphosphate |
| bp | Base pair |
| $CaCl_2$ | Calcium chloride |
| C | Celsius |
| cm | Centimeter |
| CsCl | Cesium chloride |
| CTX | Cholera toxin |
| CHEF | Clamped homogenous electric fields |
| cfu | Colony forming unit |
| DNA | Deoxyribonucleic acid |
| $dH_2O$ | Distilled water |
| DS | Drug susceptible |
| EtBr | Ethidium bromide |
| EDTA | Ethylenediaminetetraacetic acid |
| XDR-TB | Extensively drug resistant tuberculosis |
| gp | Gene product |
| g | Gram |
| hr | Hour |
| HIV | Human immunodeficiency virus |
| HCl | Hydrochloric acid |
| HMC | hydroxymethylcytosine |
| kb | Kilobase |
| kV | Kilovolts |
| LiAc | Lithium acetate |
| LA | Luria-Bertani Agar |
| LB | Luria-Bertani broth |
| LBSG | Luria-Bertani broth sucrose glycine |

| | |
|---|---|
| $MgCl_2$ | Magnesium chloride |
| mRNA | Messenger RNA |
| µg | Microgram |
| µl | Microlitre |
| mg | Milligrams |
| ml | Millilitre |
| mM | Millimolar |
| min | Minute |
| M | Molar |
| MW | Molecular weight marker |
| MTC | *M. tuberculosis* complex |
| MDR-TB | Multidrug resistance tuberculosis |
| MOI | Multiplicity of infection |
| nm | Nanometer |
| NRH | Not recognizably homologous |
| ORF | Open reading frame |
| Ocr | Overcome classical restriction |
| pfu | Plaque forming unit |
| PEG | Polyethylene glycol |
| PFGE | Pulse field gel electrophoresis |
| RE | Restriction enzyme |
| RNase | Ribonuclease |
| RNA | Ribonucleic acid |
| RNA Pol | RNA polymerase |
| RNR | Ribonucleotide reductase |
| rpm | Rotations per minute |
| sec | Second |
| SS | Single stranded |
| NaCl | Sodium chloride |
| SDS | Sodium dodecyl sulfate |
| SC | Synthetic complete media |

| | |
|---|---|
| TEM | Transmission electron microscopy |
| TB | Tuberculosis |
| TBE | Tris-borate-EDTA |
| Tris | Tris(hydroxymethyl)aminomethane |
| TS | Thymidylate synthase |
| UV | Ultraviolet |
| V | Volts |
| v | Volume |
| $H_2O$ | Water |
| w | Weight |
| WHO | World Health Organization |
| YPD | Yeast-extract peptone dextro |

# 1. Introduction

## 1.1 *Mycobacterium*

The genus *Mycobacterium* belongs to the family *Mycobactaeriaceae*, classified with Actinomycetes in the order of Actinobacteria. Mycobacteria are aerobic, acid fast/Gram positive, non-spore forming and non-motile bacteria (with the exception of *Mycobacterium marinum*). *Mycobacterium* has been the subject of much research due to the bacterium causing two serious and wide-spread infectious diseases: leprosy (*Mycobacterium leprae)* and tuberculosis (*Mycobacterium tuberculosis* in humans) as well as *Mycobacterium bovis* in cattle.

## 1.2 Tuberculosis

Tuberculosis (TB) is a major worldwide infectious disease that is concentrated in certain parts of Asia and sub-Saharan Africa (Zignol *et al* 2006) mainly caused by the bacillus *Mycobacterium tuberculosis* with some cases caused by *M. bovis* and *M. africanum* (LoBue *et al* 2010). *M. tuberculosis* was first identified as the causative agent of TB by Robert Koch in 1887, and individuals are infected by inhaling aerosol droplets that can contain as few as 1 – 3 bacilli. However, the immune system in a healthy individual, given several weeks, will most likely to control bacterial replication and spread, resulting in almost all infections becoming latent infections with only 5% of the infection developing into an active infection (Glickman and Jacobs 2001).

## 1.3 Use of Antibiotics

Antibiotics are antimicrobial compounds that have a low molecular weight and have been used widely starting in the 1940s. They target essential components of the cell machinery such as DNA and RNA synthesis, various cell proteins and the cell wall, thereby either inhibiting or killing the cell. For TB, the antibiotic treatment involves administering drugs isoniazid, rifamycin, and ethambutol for two months before continuing with

isoniazid and a rifamycin for another four. Should TB be drug susceptible, then the treatment has a 95% success rate (Koul *et al* 2011). But as both the use and misuse of antibiotics became more common and widespread, the world has seen an emergence of antibiotic resistance bacteria which includes *M. tuberculosis*.

**Table 1.1** Anti-mycobacterial drugs and their year of discovery. (Adapted from Nguyen and Thompson 2006)

| **First line antibiotics** | Year of discovery |
|---|---|
| Isoniazid | 1952 |
| Rifampin | 1966 |
| Pyrazinamide | 1952 |
| Ethambutol | 1980 |
| Rifapentine | 1965 |
| **Second line antibiotics** | |
| Cycloserine | 1952 |
| Ethionamide | 1956 |
| *P*-aminosalicylic acid | 1946 |
| Streptomycin | 1944 |
| Amikacin | 1972 |
| Kanamycin | 1957 |
| Capreomycin | 1963 |
| Levofloxacin | 1986 |
| Moxifloxacin | 1996 |
| Gatifloxacin | 1992 |

### 1.4 Antibiotic Resistance

The misuse of antibiotics has resulted with TB being a recurrent problem, especially in the developing world. Moreover, most of the first and second line antibiotics used today were discovered and introduced about forty years ago (Nguyen and Thompson 2006). With the emergence of antibiotic resistance mycobacterium, the situation is now made

much more complicated and requires a need to discover both new cell targets and antibiotics to act upon them. There are two types of resistance: The first type of resistance involves the two of the main first line drugs isoniazid and rifampicin, this degree is known as Multi-Drug Resistance (MDR-TB). The more resistant type is the Extensively Drug Resistant tuberculosis (XDR-TB) which has the same resistance of MDR-TB as well as any second line TB drugs such as kanamycin or amikacin and also any fluoroquinolone (Koul *et al* 2011).

The emergence of drug resistance is rapidly reversing the gains antibiotics have made as more new cases of MDR and XDR-TB emerge. It is estimated that there will be about 9.8 million new cases of TB in 2011, most of them in developing countries where resources are poor and the effect of health care limited (Dye and Williams 2010). With the spread of the human immunodeficiency virus (HIV), the reactivation of TB from a latent infection to that of an active one is rapidly increasing in these immunocompromised individuals. A quarter of all HIV positive patients die of TB and about 80% of these HIV-TB co-infections are located in Africa (WHO 2010). Furthermore, co-infected patients of both TB and HIV experience severe complications in their treatment regime. This is due to an increase in the amount of pills taken and thereby an increase in drug-drug interaction which leads to a decrease of overall effectiveness of drugs and could lead to health concerns in regards to side effects involved (Koul *et al* 2011).

The increase of cases of resistant TB not only increases the chances of failure of conventional treatment, but also ensures the persistence of a more virulent strain within a community and in turn increases the risk to uninfected individuals by exposing the general community to resistant strains.

### 1.5 Mechanism of Resistance

Bacteria have developed various modes of resistance that seeks to keep antibiotics out of the cell such as selective permeability of the cell wall, porin channels and efflux pumps.

Or once antibiotics have entered the cell, neutralize them with various antibiotic modifying and target modifying enzymes.

### 1.5.1 Porin Channels and Selective Impermeability

The current model of mycobacterial cell wall has a cell envelope that is supported by a peptidoglycan sacculus and a cell wall containing two layers that includes a long chain lipophilic mycolic acid bilayer that is asymmetrical and helps in limiting the entry of hydrophilic molecules (Jarlier and Nikaido 1994). Wrapped underneath it is an arabinogalactan layer that prevents hydrophobic molecules from entering (Brennan and Nikaido 1995). Anchored to this is a waxy and non fluid mycolate barrier which allows it to limit the entry of both hydrophobic and hydrophilic compounds (Liu *et al* 1995). This structure is further reinforced by a trehalose dimycolates and glycopeptidolipids network, forming the exterior. Therefore, certain enzymes such as FbpA that are responsible for mycolic acid chain assembly and construction of the envelope are also responsible in conferring various degrees of resistance to a range of antibiotics (Nguyen *et al* 2005).

### 1.5.2 Efflux Pumps

The outer envelope of cell plays an important role in conferring antibiotic resistance to Mycobacteria, considerably slowing the rate of drug entry into the cell (Jarlier and Nikaido 1994). This rate of entry allows the cell enough time to activate resistance genes to either neutralize or export these compounds, including efflux transporter genes that can be found across the genome in *M. tuberculosis*. These genes encodes for 26 ATP binding cassettes (ABC) transporters and 16 major facilitator proteins (Braibant *et al* 2000; De Rossi *et al* 2002). Research has experimentally demonstrated that transporters confer resistance to a range of antibiotics, including fluoroquinolones, rifampicin, chloramphenicol, tetracycline and aminoglycosides (Nguyen and Thompson 2006).

### 1.5.3 Antibiotic Degrading and Target Modifying Enzymes

Another form of resistance is the ability for the cell to degrade the small amounts of antibiotics that have penetrated through the cell wall. An example of this is the expression of β-lactamases, which degrades β-lactam antibiotics that undermines cell wall structure by inhibiting peptidoglycan assembly enzymes (Chambers *et al* 1995). This enzyme is encoded by the genes *blaS* and *blaE* in *M. smegmatis* and *blaC* in *M. tuberculosis* and is also produced under normal growth conditions (Flores *et al* 2005). Another strategy is to modify the target that antibiotic acts upon, causing the drug to have no effect on the cell. One example is erm37, which is found in the *M. tuberculosis* complex (MTC) and confer resistance to macrolide and lincosamide (Nguyen and Thompson 2006).

### 1.5.4 Mimicking Drug Targets

A new way of antibiotic resistance involves genes producing proteins that resemble an antibiotic target. An example of this is the gene *mfpA*, which encodes a protein that resembles a DNA double helix (Hegde *et al* 2005). This confers resistance to fluoroquinolones, which acts against mycobacteria by binding to the enzymes DNA gyrase and topoisomerase IV. This ensures that the DNA strands are not resealed, thus ensuring its eventual degradation which in turn will lead to cell death (Nguyen and Thompson 2006).

### 1.5.5 Regulation of Resistance Mechanisms

All intrinsic resistance mechanisms described above are controlled by a system that seems to have been retained in mycobacterium species and other bacterial species in the Actinomycete taxon (Morris *et al* 2005). In both antibiotic producing and pathogenic Actinomycetes (such as Streptomyces), a transcriptional regulator WhiB7 controls a set of resistance genes found in eight loci with four open reading frames (Philalay *et al* 2004) that are responsible in dealing with antibiotics that have entered the cytoplasm. A large

part of the resistance mechanism in mycobacteria is also controlled by a WhiB7 homologue thought to also have evolved in the same natural environment (Nguyen and Thompson 2006). WhiB7 not only confers resistance to various antibiotics but also plays a role in the pathogenicity of mycobacteria, thereby ensuring its long term survival during the latency period

### 1.6 Bacteriophage Mediated Antimicrobial Activity

The ease of spread of TB, the lack of new drugs being introduced and frequent misuse or overuse of these drugs, combined with a comprehensive intrinsic resistance mechanism has resulted in the steady erosion of in the ability of antibiotics to act against TB. This, coupled with the rapid spread of HIV among developing nations, especially in sub-Saharan Africa, and the mounting challenge in the development of new antiretroviral and antimicrobial drugs that are compatible with each other has called for new approaches in tackling the global TB problem.

One solution lies in finding novel antimicrobial targets and the synthesis of new antimicrobial drugs to act against them. This solution focuses on the understanding the *M. tuberculosis* genome. With the advance in sequencing technology, the first bacterial species to have its genome completely sequenced was done on *Haemophilus influenzae* in 1995 (Fleischmann *et al* 1995) paved the way for more of the same sequencing of other bacterial species and showed hundreds of genes that are vital for *in vitro* growth confirmed by mutagenesis experiments (Brown 2004). While these studies provide information on potential targets for developments of new antimicrobial drugs, it only serves to confirm their importance in maintaining cell vitality yet provides no new information as to their exact cellular function. To further understand functions of these potential targets and to prioritize them for drug discovery, one approach adopted by Lui *et al* in 2004 focused on the use of bacteriophages.

## 1.7 Phage Diversity

Bacteriophages, especially lytic phages, are absolute parasites to bacteria. Therefore they have billions of years of evolutionary history with their host and thus have ensured that phages are not only abundant in numbers in the virosphere, but also the most abundant source of genetic diversity on the planet (Suttle 2007). This phenomenon is first observed from studies done on marine phages (Bergh *et al* 1989; Suttle *et al* 1990), where bacteriophages play an important role of nutrient recycling and also transfer genetic material between host species. Bacteriophages are also abundant in numbers, with an average of 10 million viruses per milliliter of sea water (Bergh *et al* 1989) with numbers for any given phages depends on the numbers of their corresponding hosts.

Phage diversity was first observed when electron microscopy performed on marine phages which show that phages are morphologically diverse (Torella and Morita 1979). Other studies focused on the target genes (Wilson *et al* 1999, Frederickson *et al* 2003) revealed profound genotypes difference of various communities that were separated by a mere few meters, showing that marine phages consist of extremely diverse communities existing in close proximity to each other.

These studies, however, cannot provide information about overall genetic diversity of phage communities because the target genes that are used are found only in phages that can be cultured. An approach to over come this involves using metagenomic studies, where phages are directly collected from water samples and sequenced through shotgun sequence techniques or Sanger sequencing techniques. One such metagenomic study (Breitbart *et al* 2002) examined two uncultured phage communities from costal areas. The results show that over 65% of sequenced genomes do not show any sequence similarity to those in the database. Of the sequences that show various degrees of similarities, only a third of these sequences had their closet match to known phage sequences.

Such diversity is apparent in the mosaic arrangement of phage genes in their genome, by which phage genomes consist of assemblies of various genetic modules. These modules can be a part of a gene, a whole gene or a group of genes and can be found in phages that are not related. This is observed not only in sequenced marine phages but also in cultured phages as well. In a review (Casjens 2005) that examined the diversity within *Caudovirales* shows that although various tailed phage groups that infect a common host (*E. coli*) share similar lifestyles, gene arrangements, number of genes and transcription patterns, they vary greatly in their sequence similarity. An example could be seen in the sequencing of some 20 phages that are related to the phage λ and yet are all related to one another only through their mosaic arrangement of genes. Two phages, P22 and N15, have a number of genes that are very similar to those in λ but have very little sequence similarities with themselves (Hendrix and Casjens 2005; Casjens 2003).

### 1.8 Mosaicism

There are two types of mosaicism: Quantitative and qualitative. The former is the recombination of sequences from homologous regions of phage genomes, thus creating a family of homologous genes through simple divergence. This can arise from genes that are recently diverged (and are almost identical to each other) to genes that share such a long divergent history that their similarities can no longer be recognized since they have diverged beyond such point. Qualitative (or modular) mosaicism results from the non-homologous (illegitimate) recombination of genes that are either very divergent homologue or a non-homologue (also know as Not Recognizably Homologous genes – NRH). Hence when these phages are compared, it will result in patches of regions where sequence similarities between abruptly ends, resulting in regions know as novel sequence joints, with most of such joints found at gene boundaries. These illegitimate recombinations occur randomly with unnecessary genetic material deleted and nearby genes rearranged. Studies (Hendrix 2003; Pedulla *et al* 2003) have identified novel joint regions containing these genetic materials supporting this model of recombination. If the exchange of these modules occurred recently in its evolutionary time frame, then its sequence similarity will still be similar to each other. If the exchange occurred some time

ago in evolutionary time, then the only similarity would only show through from the two amino acid sequence being statistically significant.

### 1.9 Diversity within Mycobacteriophages

Mycobacteriophages also exhibit a high amount of mosaicism, indicating its high diversity. A study done by Hatful *et al* in 2010 analyzed 60 known mycobacteriophages sequences. Of the 60 sequences, 55 of them were grouped into nine clusters that share sequence similarities with the last 5 not fitting into any of the nine clusters. Of those nine clusters, five of them were able to be divided into sub-clusters with the phages within each clusters varied greatly in their sequence similarities to each other. An example of this is shown for 6 genomes within the Cluster D sharing on average 97.5% of sequence similarity while 2 genomes within Cluster I having similarities that were not apparent. In addition to comparing whole genomes, Open Reading Frames (ORFs) were also predicted for the 60 mycobacteriophages. A total of 6858 ORFs were predicted and grouped into clusters of related sequences called phamilies (phams). Close to half (46%) of these phams in a total of 1523 contains only a single ORF while 18.8% of the total phams have sequence similarity to other bacteriophages. Assigning functions to these ORFs remain a difficult as only 10% of all phams can be identified with a known function on the database. This clearly shows the great amount of variety with a small sample size (60) of mycobacteriophages and identifying this type of phage to be greatly diverse like all other phages communities so far encountered.

### 1.10 Application of Bacteriophages as Antimicrobial Agents

Given the ability of phages to not only destroy its host cell, but replicate exponentially. It has been used in various ways as an antimicrobial agent against a variety of infectious diseases. One such use is the Direct Phage Therapy. This was first pioneered by Felix d'Herelle in 1917 in which he successfully treated patients with intestinal dysentery after successfully isolating the first phage. But with the advent of antibiotics in the 1940s, phage therapy was largely abandoned in the West after World War II, but advances in the

clinical application of phages was made many advances in the Soviet Union and the Eastern Bloc due to these states being cut off from Western medical technology with the advent of the Cold War. Many phage research and production facilities were set up in the USSR both before and after WWII. These facilities produced phages that can be applied in many different forms, from orally ingested liquids and tablets to locally applied creams (Chanishvilli 2009). Soviet phage research enjoyed a high degree of success (Kutateladze and Admia 2010), but their knowledge on phages were slow to diffuse to the West.

With the growing threat of antibiotic resistance, however, phages now serves as a viable alternative as an antimicrobial agent. According to Kutateladze and Admia, when compared to antibiotics, phages offer several advantages:

1) Since phages multiply naturally, contrasting to synthetic antibiotics, they can act as a self regulating drug. Unlike an antibiotic that needs to be constantly ingested at a regular interval (pill burden) due to clearance from the body, the number of phages will naturally decrease once all bacterial cells are eliminated.
2) The restricted host range of phages means that it is only able to target specific bacterial species, unlike antibiotics that can eliminate a broad range of bacterial cells, even if they are beneficial.
3) Phages do not have harmful side effects which is generally associated with wide spectrum antibiotics.
4) Even if the bacteria show resistance to a single phage or a cocktail of them, the host must mutate or lose one or more of its receptors. This could lead to a reduction in the fitness of the cell which could help slow down the progression of disease. If the receptor is a virulence determinant, then bacterial virulence will decrease.

Despite the need to find a new alternative to antibiotics, questions are still being raised about the feasibility of phage therap. Skurnik and Strauch (2006) argued that more information is needed on the fate of phages in humans and animals, and how these phages can perform *in vivo* and also the *in vivo* data of one phage in relations to another. The

latter is important if various phages are to be administered together as a cocktail. Furthermore, the functions of vast amount of phage genes are still unknown with sequence similarity to unknown genes from various organisms. Research has shown that certain phages carry virulent factors and toxin genes. Examples include phage conferring pathogenicity to *Streptococcus pyogenes* (Broudy and Fischetti 2003), and CTX cholera (Davis and Waldor 2003) and shiga toxin (Strauch *et al* 2004). Hence, no phage therapy should be attempted until the genome of any candidate phage is sequenced and its interaction with the host fully understood.

### 1.11 Phage Products

Besides direct phage therapy, another way of using bacteriophages as antimicrobials is to use phage produced enzymes, particularly lytic enzymes (lysins) which target the bonds of the peptidoglycan of the cell wall (Young 1992). Like the phages from which they are isolated and produced, most lysins can only target specific species of bacteria (Fischetti 2005). This ensures that only the target pathogen will be lysed, leaving beneficial microorganisms within the human body untouched. However, research has shown that lysins can also be broad spectrum and target multiple bacterial species. An example of this is the lysin isolated from enterococcal phages that are able to target *Staphylococcus aureus*, and *S. pyogenes* (Yoong *et al* 2004).

So far, research has show that there is no case of reported resistance against lysins. One possible explanation is that because phages have co evolved with bacteria for millennia, it has developed strategies to prevent itself from getting trapped inside its host. By developing lysins that target essential molecules of the cell wall, it makes resistance against it very difficult (Fischetti 2005). This has the potential to make lysins the ideal candidate for an alternative to antibiotics but can also shed light on new bacterial targets for the development of new antimicrobial drugs. Lysin is but one of the many phage products that can act against bacteria.

### 1.12 Phage-Host Interactions

Other studies have been conducted to determine the genetic and molecular mechanism by which bacteriophages increases its efficiency of infection of the host by interacting with host cells at a genetic level. The interactions at this level are focused on the inhibition, activation or redirection of host proteins to phage use and such interactions occurs at all stages of the infection cycle, encompassing host transcription, translation and replication.

**Table 1.2** Review of some bacteriophage-host interactions, unknown modification are left blank (adapted from Roucourt and Lavigne)

| Phage | Phage protein | Host protein | Modification | Effect | Length | Stage |
|---|---|---|---|---|---|---|
| T7 | gp0.3 | Type 1 restriction endonuclease | | Protection of genome from restriction | 117 | Early |
| | gp0.7 C-term | RNA Pol β' subunit | | Host transcription shutoff | 117 | Early |
| | gp0.7 N-term | RNA Pol β and β' subunit | Phosphorylation | Efficiency of termination | 242 | Early |
| | | RNase III | Phosphorylation | Processing of T7 mRNA | | |
| | | RNase E and RhIB | Phosphorylation | Protection of T7 mRNA from degradation | | |
| | | Translation, initiation and elongation factors | Phosphorylation | Unknown | | |
| | gp2 | RNA Pol β' subunit | | Inhibition of transcription initiation | 64 | Middle |
| | gp1.2 | dGTPase | | Inhibition of GTPase | 85 | Early |
| T4 | Alt | One RNA Pol α subunit | Ribosylation | Preferential expression of T4 early genes | 675 | Late |
| | Alc | RNA Pol β' subunit | | Host transcription shutoff | 167 | Early |
| | ModA | Both RNA Pol α subunit | Ribosylation | Lower expression of T4 early and host genes | 200 | Early |
| | ModB | S1 ribosomal protein, trigger factor, Ef-Tu | Ribosylation | Unknown | 207 | Early |
| | AsiA | $\sigma^{70}$ factor | | Inhibition recognition of $\sigma^{70}$ promoters | 90 | Early |

| Phage | Phage protein | Host protein | Modification | Effect | Length | Stage |
|---|---|---|---|---|---|---|
| | MotA | $\sigma^{70}$ factor | | Recognition of T4 middle promoters | 211 | Early |
| | gp55 | RNA Pol core enzyme | | σ factor for expression from late promoters | 185 | Middle |
| | gp33 | RNA Pol core enzyme | | Helper of gp55 | 112 | Middle |
| | RpbA | RNA Pol core enzyme | | Favours expression of T4 late genes | 129 | Early/Middle |
| Twort | gp168 | DnaN (DNA Pol III β subunit) | | Shutoff host replication | 74 | Unknown |
| G1 | gp240 | DnaN (DNA Pol III β subunit) | | Shutoff host replication | 58 | Unknown |
| λ | P | DnaB (DNA helicase) | | Initiation of λ genome replication | 233 | Delayed Early |
| | N | RNA Pol core, NusA, NusG | | Antitermination, delay early transcription | 133 | Early |
| | Q | RNA Pol holoenzyme (core + $\sigma^{70}$) | | Antitermination, late transcription | 207 | Delayed early |

However, most of these phage proteins identified do not show sequence similarities to other known proteins and hence very little is known about their exact functional mechanism. Even within the well known and studied phage T4, about half of predicted ORFs do not have a known function, and majority of these ORFs are transcribed by early promoters (Miller 2003). But since some early phage proteins clearly exhibit antimicrobial properties, most research focus have shifted to using such proteins to identify new antimicrobial targets (Liu 2004) while ignoring other non inhibitory proteins, making most phage-host interactions remained unstudied.

### 1.12.1 Interaction in Host Defence and Patterns of Phage-Host Co-Evolution

One advantage of using bacteriophages as antimicrobial agents is the ability of phages to co-evolve with their bacterial host in order to overcome host defence mechanism against viral infections (Tock and Dryden 2005). The bacterial restriction-modification system contains two enzymatic functions that allow the host to protect itself against the entry of foreign DNA (such as those of phages). One function employs restriction endonuclease (RE) to recognize and cleave a specific DNA sequence with the other employing a methyltransferase that will modify host DNA by means of host methylation of adenine or cytosine bases. This ensures that the foreign DNA will be destroyed through cleavage by RE while the host DNA will be protected through modification of bases within the recognition sequence (Boyer 1971, Yuan 1981). However, with the co-evolution between bacteriophages and their hosts, phages have developed several strategies to overcome such defence, they include: Inhibition of RE (Tock and Dreyden 2005); degradation of enzyme cofactors; methylation of phage DNA; sequence alteration; and incorporation of hypermodified bases in phage DNA (Warren 1980).

## 1.13 Phage Mediated Target Identification

The interactions discussed above shows that bacteriophages employ sophisticated mechanisms to influence host replication machinery and metabolism. Given that most phages are extremely diverse and share very little sequence similarities with each other,

characterization of function for phage proteins remains difficult. However, although limited in understanding of the functionality of phage proteins and their interactions, several studies are now providing new information about some of the interactions discussed above. This allowed Roucourt and Lavigne (2009) to propose several hypotheses that are applicable to early phage proteins:

1) Known interactions are not isolated events but rather very common with bacterial and archeal viruses and are directed by early proteins.
2) Most interactions are not essential to infection, although they increase efficiency of infection.
3) Of these non-essential proteins, only a small fraction of these proteins are needed for efficient infection in very specific conditions
4) It is likely that any interaction between any given phage and its host is dependent on the proteins that are encoded by that particular phage or more specifically, what is not encoded by the phage. This mean phages that lack certain mechanisms or metabolic functions will encode proteins that will redirect that particular mechanism/function to increase its efficiency in the infection cycle.

Although most understanding of phage proteins and their specific individual functions remain limited, recent studies into whole genomes of bacteria has provided many in depth information on bacterial genes and the importance of certain bacterial genes to the functioning of the cell. Through many mutagenesis studies in various bacteria, results suggest a wide range of genes are needed by the cell for growth and much more for pathogenesis (Mecsas 2002). This provides a wide range of potential targets for synthesis of new antimicrobial drugs. However, these studies do not provide information of their functions, but only whether a given gene is essential to viability or pathogenesis. Since phage strategies consist of encoding its own proteins to interact with that of the host with the aim of increasing infection efficiency and eventual destruction of the host cell, continuous studies into phage proteins will provide insights into phage tactics developed over billion of years of evolution relationship between viruses and hosts and also specific functions of host genes. Hence, the results shown in these investigations into the

functionalities of bacteriophage proteins will not only have importance in phage biology, but will also have practical applications from phage therapy to more direct applications such as identifying new targets for the development of novel antimicrobial drugs that can produce similar effects to phage proteins

A study done by Liu *et al* in 2004 first used this approach to first discover antimicrobial targets before prioritizing them for drug discovery. This study focused on the pathogenic bacterium *Staphylococcus aureus* and their associated phages. A sample of 26 lytic *S. aureus* phages was sequenced and the genome analyzed through controlled expression of phage open reading frames (ORFs). Phage products that display inhibitory characteristics were selected for and these 26 phages produced 895 ORFs which yield 31 different families of inhibitory proteins. These proteins were then analyzed through affinity chromatography whereby host proteins were isolated when passed through an affinity matrix of phage inhibitory proteins. Identification of these proteins was then done through mass spectrometry, and 7 phage proteins, with size ranging from 25 to 297 amino acids were derived from such analysis. Their targets include proteins from some of the host transcription and translational machinery mentioned in Table 1.2, such as β-subunits of DNA polymerase III and helicase loader. These replication enzymes were confirmed to be indispensible to the cell when their chromosomal copies were deleted and through conditional rescue with plasmids containing these copies expressed through an IPTG inducible promoter. Since none of these host proteins in the replication machinery are currently being targeted by antibiotics (with the exception of DNA topoisomerase), the identification method employed by Liu *et al* can help extensively expand the list of essential host targets that can be exploited as new antimicrobial drugs are synthesized.

### 1.14 Aim of the Study

The aim of this study is to identify novel mycobacteriophages and their phage products which are inhibitory to *Mycobacterium smegmatis*. Phages will be isolated from various soil samples and then characterized. Genomic libraries will be built using shuttle vectors transformed into *E. coli* and individual clones obtained. These clones will then be transformed into *Rhodococcus erythropolis* to screen for antimicrobial properties. Clones that exhibit these properties will be sequenced and their similarities examined against those in the database. Whenever possible, minimum necessary DNA will also be determined on these clones. Chosen clones will also be transformed into *Saccharomyces cerevisiae* to see whether they exhibit antimicrobial properties in eukaryotes.

# 2. Material and Methods

## 2.1 Bacterial and Phage Strains used and Growth Conditions

All bacterial sstrains used were sourced from Sibayama 2006 and Sibayama 2011. All phage strains used in this study are isolated from soil samples mentioned in this work. Exception is the rifampicin and streptomycin double resistant mutant of *M. smegmatis*, which was generated from non resistant *M. smegmatis* $MC^2155$ and grown in 100µg/ml of rifampicin and streptomycin. Bacterial strains used in this work are listed in the table below and phages used throughout this work is listed in Table 2.2

**Table 2.1** Bacterial strains used in this study

| |
|---|
| *N. aobensis* IFM 10795 |
| *N. nova IFM* 10797 |
| *N. brasiliensis* IFM 10745 |
| *N. farcinica* IFM 10757 |
| *N. farcinica* IFM 10779 |
| *M. parafortuitum* 490 |
| *M. smegmatis* $MC^2$ 155 |
| *G. rubropertincta* ATC25593 |
| *R. erythropolis* 4277 |
| *R. rhodochrous* 01 |
| *R. equi* |
| *R. erythropolis* SQ1 |
| *E. coli* MM294-4 |

**Table 2.2** Phages used in this work. Nomenclature of phages was based on the location from where their respective soil samples originates. In the case of the "Botswana" phages, four phages were isolated from two soil samples collected from Botswana, with "Botswana 1" collected from Gaborone and "Botswana 2" from soil samples collected from Mochudi

| |
|---|
| Mozambique 1 |
| Mozambique 3 |
| Caltech 1 |
| Caltech 3 |
| Botswana 1 A |
| Botswana 1 B |
| Botswana 2 A |
| Botswana 2 B |

All bacterial liquid cultures were grown in Luria Broth (LB) medium agitated on a shaker or a wheel and all plate cultures are grown on solid LB medium with 1.5% agar (LA).

For bacterial growth conditions, *E. coli* MM264-4 was grown in 37$^{o}$C overnight supplemented with 100μg/ml nalidixic acid. *M. smegmatis* MC$^{2}$155 was grown in 37$^{o}$C for two days supplemented with 100μg/ml rifampicin while double resistant mutant used for transduction was grown with 100μg/ml rifampicin and streptomycin. For *R. erythropolis* SQ1, culture was supplemented with 50μg/ml of rifampicin. All other bacterial strains were grown in 30$^{o}$C for 2 – 4 days. All bacterial cultures were stored at 4$^{o}$C and for long term storage; cultures were stored in -70$^{o}$C in 33% glycerol.

Liquid *S. cerevisiae* culture was grown in YPD medium at 30$^{o}$C agitated on a shaker.
All phages were propagated with *M. smegmatis* MC$^{2}$155. For transduction, phages were propagated with its rifampicin and streptomycin double resistant mutant. All phage suspensions were stored at 4$^{o}$C.

All media, reagents and their composition can be found in Appendix A.

## 2.2 Plasmids

All plasmids used throughout this work are listed in the table below

**Table 2.3** Plasmids used in throughout this work

| Plasmid | Characteristics | Source |
|---|---|---|
| pDA71 | *E. coli* - *Rhodococcus* shuttle vector; Amp-R | Y. Shibayama 2006, 2011 |
| pUC 18/19 | *E. coli* cloning vector; LacZ'; Amp-R | Fermentas |
| pSHLeu | *S. cerevisiae* cloning vector; Kan-R; Leu | M. Kitakawa (X. Gan *et al* 2002) |

In *E. coli*, transformants for both pUC18/19 and pDA71 was selected with 100µg/ml ampicillin and in *R. erythropolis*; 40µg/ml was used in the selection for transformants. For pSHLeu, 100µg/ml kanamycin was used to select for transformants in *E. coli*, while for *S. cerevisiae*. Plasmids were stored in 6M CsCl.

## 2.3 Isolation and Propagation of Mycobacteriophages

### 2.3.1 Soil Assay

One gram of soil sample in 10ml LB supplemented with 10mM of $CaCl_2$ and 10mM $MgCl_2$ (thereafter referred to as divalent ions and for all subsequent supplementation of both these ions the concentration remains 10mM) was agitated on a shaker for 16hrs in 37$^{o}$C at 90 – 100 rpm. This soil suspension was then centrifuged in a Beckman JA-20 rotor at 15000 rpm for 10min. 2ml of supernatant was added to 30µl of *M. smegmatis* stationary culture in 1ml of LB in a test tube. This mixture was then briefly vortexed for 1sec and then incubated at 37$^{o}$C for 30min. 3ml of sloppy agar was then added to this mixture, mixed by rolling the test tube between the palms and poured onto a LA plate supplemented with 100µg/ml rifampicin and 50µg/ml nystatin, and divalent ions. The overlaid plate was then incubated at 37$^{o}$C for 2 days and observed for plaques.

### 2.3.2 Production of Phage Suspension and Single Plaque Purification

A single plaque was picked with a sterile toothpick piecing through the sloppy agar and mixed into 100µl of $dH_2O$. This was then briefly vortexed, left to stand on bench for 10min and microfuged for 2min. The phage containing supernatant was then serially diluted to determine the titre of the phage suspension and also the dilution required to produce single plaques. 10µl of the diluted phage was mixed with 30µl of *M. smegmatis* stationary phase culture in 2ml LB in a test tube, vortexed briefly and incubated for 30 min at 37$^{o}$C. 2ml of sloppy agar was added and mixed by rolling test tube between the palms and poured onto an agar plate supplemented with divalent ions. The overlaid plate was incubated for two days at 37$^{o}$C.

## 2.4 Phage Characterization

### 2.4.1 Calculating Plaque Forming Unit (pfu)

An LA plate, supplemented with divalent ions was first overlaid with *M. smegmatis* and dried with lid open at 40$^{o}$C for 20min. A serial dilution (to $10^{-7}$) of the phage was then prepared with $dH_2O$ and10µl of each dilution was spotted onto the plate, incubated for 2 days at 37$^{o}$C or until plaques are visible. The following formula was used to calculate pfu/ml:

$$\text{pfu/ml} = \frac{\text{number of plaques} \times \text{dilution factor}}{\text{volume}}$$

### 2.4.2 Calculating Colony Forming Unit (cfu)

*M. smegmatis* culture was serially diluted to $10^{-7}$ with $dH_2O$ and 100µl of the dilutions $10^{-4}$ – $10^{-7}$ was each spread across an LA plate, incubated at 37$^{o}$C for 2 days or until colonies were visible. The following formula was used to calculate cfu/ml:

$$\text{cfu/ml} = \frac{\text{number of colonies} \times \text{dilution factor}}{\text{volume}}$$

### 2.4.3 Divalent Ion Requirements

Four LA plates: One supplemented with divalent ions; two plates, one supplemented with 10mM $MgCl_2$ only and another with 10mM of $CaCl_2$ only; and one with no ions added were prepared. Onto each plate was overlaid 20µl of stationary phase *M. smegmatis* and dried with lid open at 40$^{o}$C for 20min. Onto each plate, 10µl of phage suspension was spotted onto the plate and incubated at 37$^{o}$C for 2 days or until plaques were visible.

### 2.4.4 Host Range Test

Test strain was overlaid onto an LA plate supplemented with divalent ions and dried with lid open at 40$^{o}$C for 20min. 10µl of phage suspension was spotted onto the plate and incubated for 2 days and observed for plaques.

### 2.4.5 Transduction

Rifampicin and streptomycin double mutant prepared by spreading 300µl of rifampicin resistant *M. smegmatis* stationary culture on an LA plate supplemented with 100µg/ml rifampicin and streptomycin and incubated for 2 days or until colonies were observed. A single colony was picked with a toothpick and grown in liquid culture supplemented with 100µg/ml rifampicin and streptomycin. This liquid culture is then used for the production of phage plate lysates. Lysates was prepared, collected and titrated.

Phages and cells were mixed together to achieve two MOIs of 0.1 and 1.0 and made to a final volume of 300µl supplemented with divalent ions. Tube was inverted to mix and left to stand in room temperature for 30min. Cells were collected by microfuging for 5min and pellet was washed once with $dH_2O$ before being resuspended in 500µl of LB and spread on a LA plate supplemented with 100µg/ml rifampicin and streptomycin. Plate was incubated for 14 days and observed every 2 days for formation of colonies.

## 2.5 Lysate Production

### 2.5.1 Plate Lysate Production

A serial dilution of $10^{-1}$ and $10^{-2}$ of the phage suspension was made and 10µl of this dilution was added to 30µl of *M. smegmatis* in 2ml of LB and incubated for 30min at 37$^{o}$C. 2ml of sloppy agar was added to the plate lysate, mixed by rolling between the palms and poured onto a LA plate supplemented with divalent ions. Plates were incubated for 2 days at 37$^{o}$C or until there is a significant difference between the plate lysate and that of the cell only control.

Onto the plate lysate, 2ml of LB was added and using a spatula, all liquids and sloppy agar was gathered into a centrifuge tube. This was briefly vortexed, left for 10min and centrifuged for 10min at 15000 rpm with a JA-20 rotor. Supernatant was kept and titrated to determine whether the plate lysate has the desired titre of $\leq 10^{9}$ pfu/ml.

### 2.5.2 Production of Small Scale Liquid Lysate

Into three 100ml flasks, 200µl of stationary phase *M. smegmatis* culture was added 9.8ml of LB supplemented with divalent ions to make up total volume of 10ml and mixed by gentle swirling. Into the first flask, 100µl of plate lysate was added, mixed by gentle swirling and from the first flask, 100µl of this mixture was transferred to the second flask. This was repeated from the second flask to the third. All three flasks were incubated for 16 – 18 hrs at 37$^{o}$C agitated at 100rpm at a rotary shaker. Flasks were then collected and contents centrifuged at 15000 rpm for 10min with a JA-20 rotor. Supernatant was kept and titrated to determine whether the lysate has achieved titre of $\leq 10^{9}$ pfu/ml. Other MOIs and conditions were also tested.

### 2.5.3 Production of Large Scale Liquid Lysate

Conditions and MOIs used in the production of small scale lysates was applied in the production of 500ml large scale liquid lysate (5 × 100ml in 5000 ml flask).

## 2.6 Phage Purification in an Equilibrium Gradient

Into the large scale lysate, NaCl and polyethylene glycol (PEG) 6000 was added to a final concentration of 1M and 10% (w/v) respectively. Both are dissolved by the slow stirring of a magnetic stirrer. The lysate was then cooled in a ice slurry for 1 – 1.5hrs and centrifuged at 4$^{o}$C at 9000rpm for 20min. Supernatant was discarded and pellet resuspended in 5ml of SC buffer containing 81.7% (w/v) CsCl and loaded into a Beckman Quick-Seal centrifuge tube and centrifuged at 38000rpm at 4$^{o}$C for 24hrs. A blue purple band was observed and was extracted with a needle and syringe and titrated.

## 2.7 Transmission Electron Microscopy

### 2.7.1 Preparation of Formvar-Carbon Coated Copper Grids

A microscope slide was dipped vertically into 0.50% Formvar solution until half to three quarters submerged. Slide is removed and dried by left standing vertically for 2 – 3 min and blotted on paper towel. A scalpel was used to scrape off the Formvar on the sides and corners of the slides and was dipped slowly into a tub of $dH_2O$ at a 45$^{o}$ angle until a transparent film of Formvar was seen floating on the surface. Hexagonal 200-mesh copper grids were placed on the Formvar surface with the dull side of the grid facing downwards by means of fine forceps. The entire Formvar surface, with grids attached, was lifted onto a filter paper and air-dried on bench. The grids were then taken to be carbon coated by the Electron Microscopy Unit of the University of the Witwatersrand.

### 2.7.2 Negative Staining of Phages

Both purified phages and phages from lysates were negatively stained. For purified phages, it was diluted to $10^{10}$ pfu/ml while that of lysates were not diluted. 5µl of phage suspension was placed on the Formvar-carbon coated side of the grid and left to stand for 30min. Excess liquid was then drawn off with a paper towel and 5µl of $dH_2O$ was placed on the grid and immediately drawn off with a paper towel. 5µl of 1% uranyl acetate was placed on the grid, left to stain for 20sec before being drawn off with paper towel. Grids were left to stand for 10 – 15min to be air-dried.

### 2.7.3 Viewing of Negative Stained Phages

Negatively stained phages was viewed and photographed with a Joel JEM-100S transmission electron microscope at 80 kV. Micrographs were developed by the Electron Microscopy Unit of the University of the Witwatersrand.

## 2.8 Phage Genomic DNA Preparation

### 2.8.1 Dialysis of Purified Phages

Dialysis tubing of 5cm was boiled in 500ml of 2% $NaHCO_3$ and 1mM EDTA (pH 8) for 10min and rinsed in $dH_2O$. Tubing was then boiled in 500ml of 1mM EDTA (pH 8) for 10 min and rinsed in $dH_2O$. 500µl of purified phage suspension was pipetted into the tubing and dialyzed in 500ml of dialysis buffer of 10 mM NaCl, 50 mM Tris-HCl (pH 8.0), and 10 mM $CaCl_2$ for 1hr while gently stirring with a magnetic stirrer. This was repeated again for another hour.

### 2.8.2 Phage Genomic DNA Extraction

EDTA (pH 8), proteinase K, and SDS was added to the dialyzed phage suspension in a concentration of 20mM, 50µg/ml and 0.5% (w/v) respectively. This mixture was briefly vortexed and incubated at 56$^{o}$C for 1hr and cooled in room temperature. Three phenol extraction and one chloroform extraction was preformed on the suspension and genomic DNA precipitated using salt-ethanol precipitation, and resuspended in $dH_2O$.

### 2.8.3 *E. coli* Mini Plasmid Preparation

Single plasmid containing colony was inoculated in 1ml of LB supplemented with 100µl ampicillin and grown on a shaker for < 6hrs – overnight. Cells were pelleted by microfuging for 30sec and the supernatant decanted. 80µl of solution I was used to resuspend the cells and 160µl of solution II was added and mixed by slowly inverting several times and left to stand in room temperature for 10min. 120µl of solution III was added and mixed by shaking vigorously before being left to stand on ice for 5min. The mixture was then microfuged for 5min at 4$^{o}$C and the supernatant was kept in a new microfuge tube and warmed at 42$^{o}$C for 3 – 5 min. 220µl of isopropanol was added, inverted to mix and stood at room temperature for 5min. Plasmids was pelleted by microfuging for 5min and the pellet washed with 150µl of ethanol before being pelleted again by microfuging for 1.5min. Pellet was vacuum desiccated for 20min and plasmid DNA resuspended in 40µl of $dH_2O$ with 10µg/ml of RNase.

## 2.9 DNA Manipulation

### 2.9.1 Phenol-Chloroform Extraction

TE buffer was added to a final volume of 300µl for any DNA suspension below this volume. 100µl of phenol was added, mixed by inversion and microfuged at 4$^{o}$C for 10min, and the top layer (aqueous phase) was collected. One-third volume of chloroform was added to the collected solution, mixed by inversion and microfuged at room temperature

for 1min. The aqueous phase was again collected and subjected to salt-ethanol precipitation.

### 2.9.2 Salt Ethanol Precipitation

One tenth volume of 1M NaCl and two volumes of 96% ethanol were added to the DNA suspension and kept in ice for 30min. DNA was then pelleted by microfuging at 4$^{o}$C for 20min and vacuum desiccated for 20min. Pellet was resuspended in 100µl, or an appropriate volume of $dH_2O$.

### 2.9.3 Restriction Endonuclease Digest

Restriction endonuclease was obtained from various commercial suppliers such as Fermentas, Roche or New England Biolabs and each used according to their instructions. For single digestions, 10× digestion buffer to the DNA suspension, along with the appropriate amount of endonuclease, mixed by lightly tapping the microfuge tube and briefly microfuged. For double digestions, a common buffer which allows for performance of both endonucleases was chosen and half volume of each of the two enzymes was added, mixed and microfuged. All digestions were incubated at the instructed temperature of ≤ 2hrs.

Enzymatic activities were stopped by heat treatment according to instructions and conditions provided. If heat inactivation is not a viable option, then phenol-chloroform extraction was performed.

### 2.9.4 Dephosphorylation of Nucleic Acids

To prevent recircularization of the plasmid, Fermentas alkaline phosphatase was used. Appropriate amount of 10× phosphatase buffer and alkaline phosphatase was added and incubated for 10min at 37$^{o}$C. Enzymatic activity was then heat inactivated by incubating for 5min at 75$^{o}$C.

### 2.9.5 Ligation

Appropriate ratio of insert DNA and plasmid was added to a final volume of 17µl with $dH_2O$. 2µl of 10× ligase buffer was added along with 1µl of Fermentas ligase, tapped to mix, briefly microfuged and incubated overnight at 22$^oC$. Ligase was heat inactivated at 65$^oC$ for 10min.

### 2.9.6 DNA Agarose Gel Extraction

Location of the DNA of interest was determined by examining agarose gel under UV light and excised using a scalpel. Excised piece was crushed using a spatula, frozen at -70$^oC$ for 2hrs, thawed at room temperature and microfuged for 20min. Supernatant was kept. Tris was added to the supernatant to a final concentration of 50mM and subjected to two rounds of phenol extraction and one round of chloroform extraction. DNA was then precipitated using salt-ethanol precipitation and resuspended in 50µl of $dH_2O$.

### 2.9.7 Determining DNA Concentration

5µl of DNA sample was ran on a gel, viewed and quantified using Bio-Rad Quantity One software. DNA was quantified by comparing the known quantity of band in molecular weight marker, express in its intensity, and comparing it to the intensity of the band of the sample.

## 2.10 Gel Electrophoresis

### 2.10.1 Agarose Gel Electrophoresis

Agarose gels were prepared with concentration from 0.6% - 1.0% agarose depending on the size of DNA fragments that were separated. Appropriate amount of agarose was added to 40ml of 0.5× TBE buffer and dissolved by heating. Ethidium Bromide (EtBr) to added to the agarose solution to a final concentration of 1µg/ml and poured into a comb inserted gel

tray at 4$^{o}$C and polymerized for 20min. Comb was removed and gel tray placed into an electrophoresis unit containing 1µg/ml EtBr. DNA samples was mixed with 1 - 2µl of bromophenol blue tracking dye and electrophoresis was ran at 80 – 100V in room temperature. Gels were viewed and photographed using the Bio-Rad Molecular Imager Gel Doc XR System. Fermentas GeneRuler DNA ladder was used.

### 2.10.2 Pulse Field Gel Electrophoresis (PFGE)

A 1% agarose gel was prepared by dissolving agarose in 100ml of 0.5× TBE through heating and poured into poured into a PFGE gel tray and polymerized for 30min at 4$^{o}$C. The New England Biolabs PFG Low Range marker, embedded in agarose, was cut to a length of 1mm and placed inside a well, with remaining spaces sealed with 1% agarose in 0.5× TBE buffer and be allowed to be polymerized at room temperature for 10min. The gel was then placed in the electrophoresis unit, which was the BioRad CHEF-DR II Pulsed Field Electrophoresis System, which contains 2000ml of 0.5× TBE. DNA samples were mixed with 1 - 2µl of bromophenol blue tracking dye and loaded into the wells. Electrophoresis was conducted with initial and final switch times of 1 and 12sec respectively and at 6V/cm for 18.5hrs at 4$^{o}$C. The circular pump was switched on after 15min and ran at a circulatory rate of 750ml/min. Gel was then stained in 1µg/ml EtBr and viewed.

## 2.11 Transformation

### 2.11.1 *E. coli* Calcium Chloride Transformation

An overnight *E. coli* pre-culture was inoculated into 20ml of LB supplemented with 0.5% glucose for at 37$^{o}$C for 1.75hrs with vigorous agitation on a shaker. The new culture was cooled in a ice slurry for 5min whilst shaking, before being centrifuged in a pre-cooled JA20 rotor at 10000rpm for 10min. Supernatant was decanted and cells resuspended in 10ml $CaCl_2$ transformation buffer and placed in ice for 15min before being centrifuged at

10000rpm for 10min. Supernatant was again decanted and cells resuspended in 1.33ml of $CaCl_2$ transformation buffer and placed in ice for 2hrs.

20µl of plasmid suspension was added to 100µl of competent cells and mixed by bubbling before left to stand on ice for 15min. Cells were heat shocked at 42$^{o}$C for 90sec and 0.5ml of pre-warmed LB added and incubated at 37$^{o}$C for 1hr. Cells were spread on an LA plate supplemented with 100µg/ml ampicillin and incubated overnight at 37$^{o}$C. For vector pSHLeu, LA plate was supplemented with 100µg/ml kanamycin.

### 2.11.2 *Rhodococcus* PEG Mediated Transformation

A *R. erythropolis* culture was grown in 5ml of LBSG (2% glycine) for 2 days in room temperature. 1ml of cells was then washed in B buffer, microfuged, resuspended again in 1ml of B buffer containing 5mg lysozyme and incubated at 37$^{o}$C for 1.5hrs with inversions every 10min. During the last 20min of incubation, 2ml P buffer was prepared and P-PEG buffer was made by preparing 0.5g of PEG6000 was UV-sterilized and dissolved in 1ml of P buffer.

Cells was microfuged for 2min, washed once with 1ml B buffer, microfuged again, and resuspended in 500ml P buffer. 5µl of plasmid was mixed with 25µl of cell protoplast by bubbling and left to stand in room temperature for 10min. 30µl of P-PEG buffer was added, mixed by bubbling, spread onto a regeneration media and incubated at 30$^{o}$C for 12hrs. After which the plate was underlay with 250µl of 4µg/ml chloramphenicol, and further incubated for 7 days or until colonies appear.

### 2.11.3 *S. cerevisiae* Lithium Acetate Transformation

1.5ml of YPD medium was inoculated with 30µl of an overnight culture of *S. cerevisiae* and incubated at 30$^{o}$C for 4hrs or until the culture reaches to a concentration of ~$2\times10^{7}$ cfu/ml. Cells was microfuged, washed once with $dH_2O$, resuspended in 250µl of 100mM lithium acetate (LiAc) and incubated at 30$^{o}$C for 10min. Cells was then pelleted,

supernatant decanted and reagents were added in the following order: 1) 25µl $dH_2O$; 2) 18µl 1.0M LiAc; 3) 12.5µl 2mg/ml SS-DNA; 4) 2µg plasmid; 5) 120µl 50% (w/v) filter sterilized PEG4000. Cell pellet was resuspended, reagents mixed by vigorous vortexing for 2min and incubated at 30$^{o}$C for 30min. Cells was then heat shocked at 42$^{o}$C for 15min, collected by microfuging for 5min, resuspended in 0.5ml of $dH_2O$, spread onto an SC plate and incubated in 30$^{o}$C for 2 days.

**2.12 DNA Sequencing and Analysis**

All DNA sequencing reactions were done by Inqaba Biotech. All sequencing DNA was prepared through *E. coli* mini plasmid preparation with a phenol extraction preformed prior to isopropanol precipitation. Sequenced DNA were analysed and edited using the following internet programs

**Table 2.4** Internet programs used for sequence analysis

| **Program** | **Website Address** |
|---|---|
| BLAST | http://blast.ncbi.nlm.nih.gov/Blast.cgi |
| FASTA | http://www.ebi.ac.uk/Tools/sss/fasta/ |
| Softberry | http://linux1.softberry.com/berry.phtml |
| GeneMark | http://exon.gatech.edu/ |
| NEB cutter V 2.0 | http://tools.neb.com/NEBcutter2/ |

**2.13 Determining Phage Novelty**

Phages used in this study had their respective genome extracted, fragmented, cloned into *E. coli* and sequenced by Iqaba Biotech. These sequenced fragments were then compared with those in the data base for any similarities using BLAST. Although some phages used in this study do have high similarities with phages in the database, they are regarded as novel in the sense that these phages do not share exact identity with those in the database. These BLAST results were not included since it would lengthen the dissertation considerably.

# 3. Results and Discussions

## 3.1 Isolation and Characterization of Mycobacteriophages

### 3.1.1 Isolation of Mycobacteriophages from Soil

To isolate mycobacteriophages, 1g of soil sample was agitated for 16hrs at 37$^{o}$C, after which the samples was centrifuged. The supernatant collected was then mixed with *M. smegmatis* MC$^{2}$155 and overlaid on a LCM plate. *M. smegmatis* was used since it is both a fast growing and a non-pathogenic species within *Mycobacterium*. This allowed for all isolation and any subsequent characterization of phages to be carried out in a safe condition in the shortest possible time. Four soil samples were collected and plaque yields are tabled below.

**Table 3.1** Soil samples and their respective phage yields

| **Soil sample** | **Plaques** |
|---|---|
| Bilene, Mozambique | 70 |
| California Institute of Technology, California, United States | 220 |
| Gaborone, Botswana (Botswana1) | 20 |
| Mochudi, Botswana (Botswana 2) | 8 |

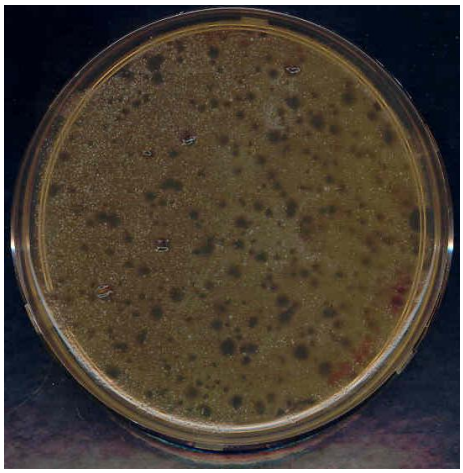


**Figure 3.1** Phage assayed from Caltech soil

Plaques observed varied in sizes and were picked using sterile toothpicks and re-suspended in 100µl of water. Phage suspensions were spotted on LCM plates overlaid with *M. smegmatis* to start single plaque purification. This was repeated three times to ensure the purity of the phage. A total of 8 phages from the above 4 soil samples were selected and purified.

### 3.1.2 Divalent Ion Requirement

Selected phages were spotted onto a lawn of *M. smegmatis* overlaid on LA plates supplemented with both divalent ions, either one of the ions or none at all. These tests whether these phages can plaque without the divalent ions calcium and magnesium. These procedures were repeated three times to ensure that these phages, which were isolated using divalent ions, did not carry through the ions in this experiment.

**Table 3.2** The divalent ion requirement of selected phages

| **Phage** | **Divalent ion requirements** |
|---|---|
| Mozambique 1 | None |
| Mozambique 3 | None |
| Caltech 1 | None |
| Caltech 5 | None |
| Botswana 1A | None |
| Botswana 1B | None |
| Botswana 2A | Calcium |
| Botswana 2B | Calcium |

### 3.1.3 Host Range

Isolated phages were spotted onto a lawn of 10 different bacterial species and incubated for 24hrs at 30/37$^{o}$C depending on the host. A table of host species used is provided below

**Table 3.3** Bacterial species used in testing host range

| | |
|---|---|
| *N. aobensis* IFM 10795 | *M. parafortuitum* 490 |
| *N. nova IFM* 10797 | *G. rubropertincta* ATC25593 |
| *N. brasiliensis* IFM 10745 | *R. erythropolis* 4277 |
| *N. farcinica* IFM 10757 | *R. rhodochrous* 01 |
| *N. farcinica* IFM 10779 | *R. equi* |

Phages Mozambique 1, Mozambique 3 and Botswana 1A each showed slight clearing on one of the 10 bacterial species tested. Mozambique 1 showed clearing on *N. brasiliensis* and Mozambique 3 and Botswana 1A showed clearing on *N. nova* and *N. aobensis* respectively. The rest of the phages showed no visible signs of plaquing or clearing.

Although none of the phages shows clear plaquing on the host strains tested, only one *Mycobacterium* species was tested. Research (Rybniker *et al* 2006) has show that different mycobacteriophages have different host range. These vary from broad activity of phage D29, which can infect both *M. smegmatis* and *M. tuberculosis*, to phage Barnyard, which can only infect a single known host. Some mycobacteriophage will even differentiate between strains of the same species.

Although many aspects of these infection barriers are not known, it is presumed that specific host cell receptors are strong factors in host preferences of mycobacteriophages, although other cellular factors such as restriction and other metabolic requirements after phage DNA injection also play a role. Mycobacteriophages such as D4 have shown to be inhibited by lipid extracts of *M. smegmatis* (Tokunaga *et al* 1970) while other molecules such as methylated rhamnose residues associated with *M. smegmatis* cell wall have been show to be involved with the infection of phage I3 (Chen *et al* 2009).

.

### 3.2 Production of Phage Lysates and Phage Purification

For the extraction of phage DNA, observation of phage morphology with TEM and to perform transduction, it was necessary to produce a large volume of phage lysate that is high in titre. The plate lysate first produced yielded a titre with an average of $10^6$ in 5mls and was subsequently used in the production of liquid lysate of 10ml which yielded a titre with an average of $10^9$. Thereafter, liquid lysates were up scaled to 500ml in preparation of precipitating these phages for DNA extraction.

#### 3.2.1 Plate Lysate

These were produced by incubating *M. smegmatis* with phage for 30 minutes at 37$^o$C in liquid medium before overlaying it on LCM agar plates and incubating them again at 37$^o$C for 48hrs. The amount of bacteria used for each case was 9 ×$10^6$. The exact concentrations of both phages used and titres are provided below.

**Table 3.4** Titres used in the production of plate lysates and subsequent titres

| **Phage** | **Initial titre (pfu) ×$10^4$** | **MOI** | **Final titre (pfu/ml) ×$10^8$** |
|---|---|---|---|
| Mozambique 1 | 20 | 0.02 | 20 |
| Mozambique 3 | 10 | 0.01 | 9 |
| Caltech 1 | 8 | 0.009 | 40 |
| Caltech 5 | 30 | 0.03 | 70 |
| Botswana 1A | 40 | 0.04 | 70 |
| Botswana 1B | 30 | 0.03 | 80 |
| Botswana 2A | 6 | 0.006 | 300 |
| Botswana 2B | 7 | 0.007 | 80 |

### 3.2.2 Liquid Lysates

Liquid lysates were produced by incubating *M. smegmatis* with phage in a 10ml broth supplemented with divalent ions calcium and magnesium and agitated on a shaker for 16hrs at 37$^{o}$C. Below are the concentration of both bacteria and phage.

**Table 3.5** Amount of phage and bacteria used in the production of small scale liquid lysates and final titres

| **Phage** | **Bacteria (cfu) $\times 10^7$** | **Initial titre (pfu) $\times 10^6$** | **MOI** | **Final titre (pfu/ml) $\times 10^9$** |
|---|---|---|---|---|
| Mozambique 1 | 15 | 20 | 0.07 | 20 |
| Mozambique 3 | 9 | 3 | 0.03 | 5 |
| Caltech 1 | 60 | 20 | 0.03 | 1 |
| Caltech 5 | 60 | 20 | 0.03 | 20 |
| Botswana 1A | 90 | 70 | 0.08 | 7 |
| Botswana 1B | 90 | 80 | 0.09 | 4 |
| Botswana 2A | 30 | 3 | 0.01 | 10 |
| Botswana 2B | 9 | 8 | 0.09 | 1.4 |

The small scale liquid lysate was in turn used in the production of large scale liquid lysate of 500ml. MOIs for the large scale liquid lysate were either kept as close to the previous MOI as possible. The amount of bacteria and phage used are listed below

**Table 3.6** Amount of phage and bacteria used in the production of large scale liquid lysates and final titres

| **Phage** | **Bacteria (cfu) $\times 10^7$** | **Initial titre (pfu) $\times 10^6$** | **MOI** | **Final titre (cfu/ml) $\times 10^9$** |
|---|---|---|---|---|
| Mozambique 1 | 15 | 20 | 0.07 | 3 |
| Mozambique 3 | 3 | 1 | 0.03 | 5 |
| Caltech 1 | 3 | 1 | 0.03 | 2 |
| Caltech 5 | 6 | 2 | 0.03 | 10 |
| Botswana 1A | 9 | 7 | 0.08 | 6 |
| Botswana 1B | 4.5 | 4 | 0.09 | 10 |
| Botswana 2A | 10 | 1 | 0.01 | 7 |
| Botswana 2B | 1.5 | 1.4 | 0.09 | 10 |

### 3.2.3 Phage Purification

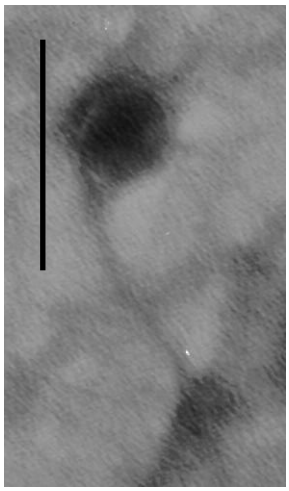
Large scale phage lysates were centrifuged and phages in the supernatant were precipitated by adding NaCl and PEG. Precipitated phages was then collected in the pellet after centrifugation and re-suspended in a smaller volume of CS buffer. Phages were then purified by a CsCl equilibrium gradient of 1.50g/ml. All phages appeared as purple blue band located slightly above or below the middle of the tube, indicating that all phages have a density range of 1.40 – 1.60 g/ml. Purified phages were titred and compared with those of the large scale lysate to calculate the percentage of recovery:

**Table 3.7** Titres of both large scale phage lysates and the subsequent purification with recovery percentage

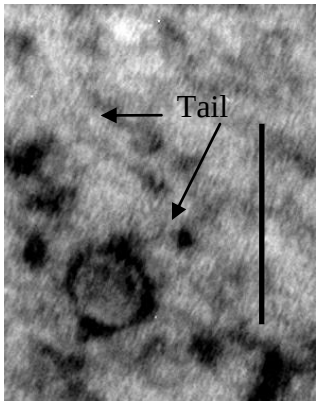
| **Phage** | **Titres of large scale lysate (pfu/ml) × $10^9$** | **Titres of purified phages (pfu/ml) × $10^{11}$** | **Recovery (%)** |
|---|---|---|---|
| Mozambique 1 | 3 | 3 | 20 |
| Mozambique 3 | 5 | 71 | 28 |
| Caltech 1 | 10 | 9 | 18 |
| Caltech 5 | 10 | 4.3 | 9 |
| Botswana 1A | 6 | 3.3 | 11 |
| Botswana 1B | 10 | 2.9 | 6 |
| Botswana 2A | 7 | 2.3 | 7 |
| Botswana 2B | 10 | 7 | 7 |

### 3.3 Phage Morphology

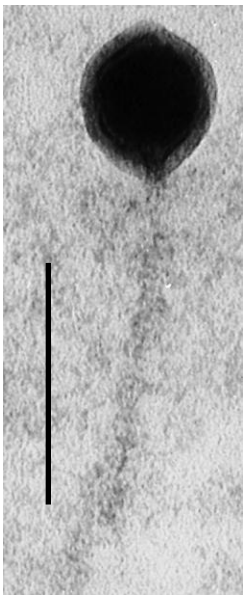
Phages were negatively stained and viewed under a transmission electron microscope, the family of each individual phage, its head diameter, and its tail length and diameter is given in the table below



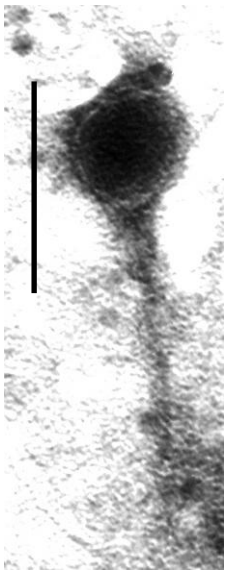
Mozambique 1




Mozambique 3

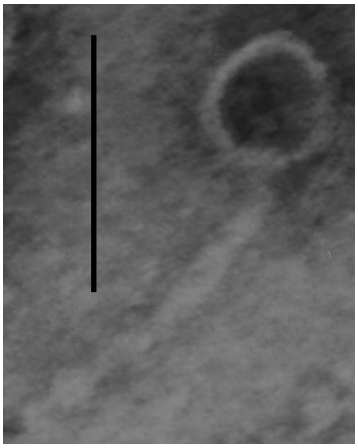


Caltech 1

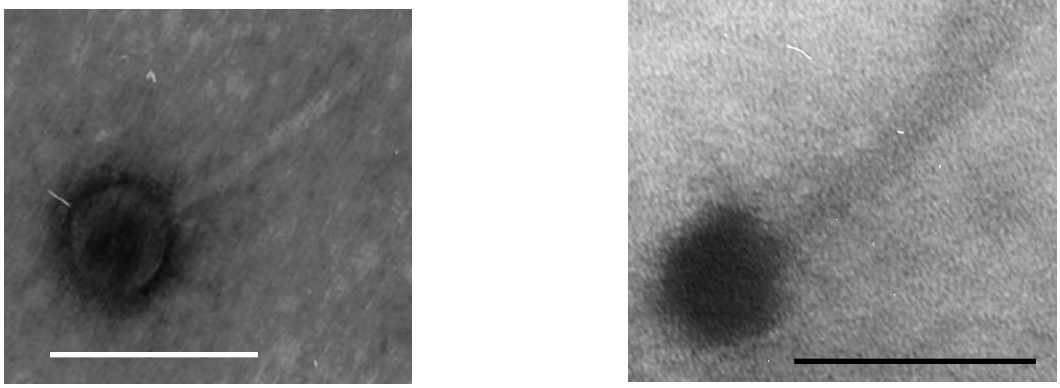


Caltech 5

Botswana 1A



Botswana 1B



Botswana 2A Botswana 2B

**Figure 3.2** Electron micrographs of phages

Bar = 100nm. All micrographs are taken at 100 000× magnification, with the exception of phages Mozambique 3 and Botswana 2A, which were taken at 50 000× magnification.

All phages observed are found to belong to the family *Siphoviridae*. Phage dimensions were measured and are tabulated below:

**Table 3.8** Summary of phage sizes

| **Phage** | **Head diameter (nm)** | **Head length (nm)** | **Tail length (nm)** |
|---|---|---|---|
| Mozambique 1 | 54 ± 3 | 56 ± 3 | 116 ± 3 |
| Mozambique 3 | 60 ± 0 | 60 ± 0 | 136 ± 24 |
| Caltech 1 | 52 ± 2 | 60 ± 0 | 150 ± 0 |
| Caltech 5 | 54 ± 3 | 56 ± 8 | 148 ± 7 |
| Botswana 1 A | 54 ± 3 | 54 ± 3 | 114 ± 3 |
| Botswana 1 B | 56 ± 3 | 58 ± 2 | 124 ± 3 |
| Botswana 2 A | 60 ± 0 | 68 ± 6 | 124 ± 4 |
| Botswana 2 B | 56 ± 3 | 60 ± 2 | 146 ± 3 |

The values above are the mean values obtained from measurements done on five different images of phage particle per phage. Statistical tests were performed using the following formula:

$$s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

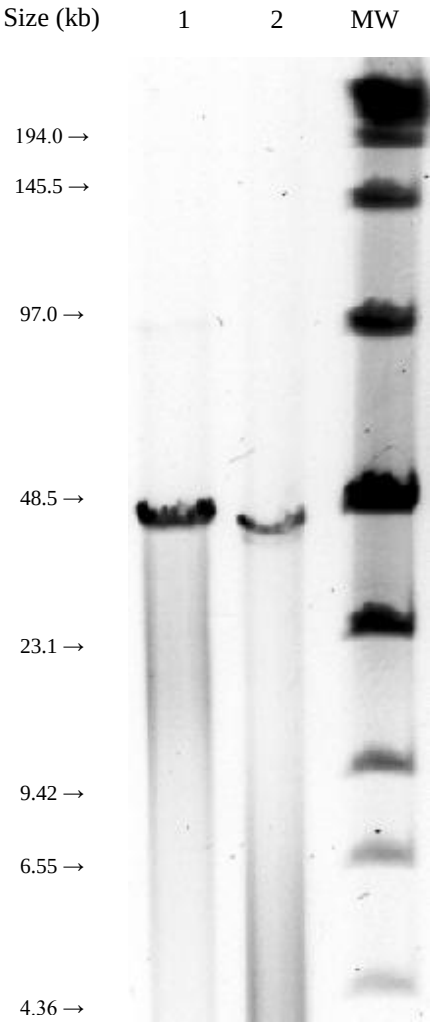
$S^2$ is the unbiased estimator of the samples. N is the sample size and X is any given measurement.

In a review (Ackermann 2007) which examined the morphology of 5500 phages, it was established that almost all phages (96%) are tailed phages, and 61% of these phages belong to the family *Siphoviridae*. Within mycobacteriophages, all phages so far characterized have been double stranded DNA tailed phages and in a recent review (Hatfull 2010) of 70 mycobacteriophages, belong to the family *Siphoviridae*, which are characterized by long non-contractile tails varying from ~105nm to ~300nm. The review also found that head shapes and sizes vary widely amongst these phages. Although most heads are isometric, three of the 61 phages contain prolate heads, with some heads having a length to width ratio of four. Ranges of head sizes range from the smallest of 48nm in diameter for phages BP and Halo to the largest of 85nm of phage Bxz1. Other than head sizes and shapes, structure of tail tips also vary from phage to phage.

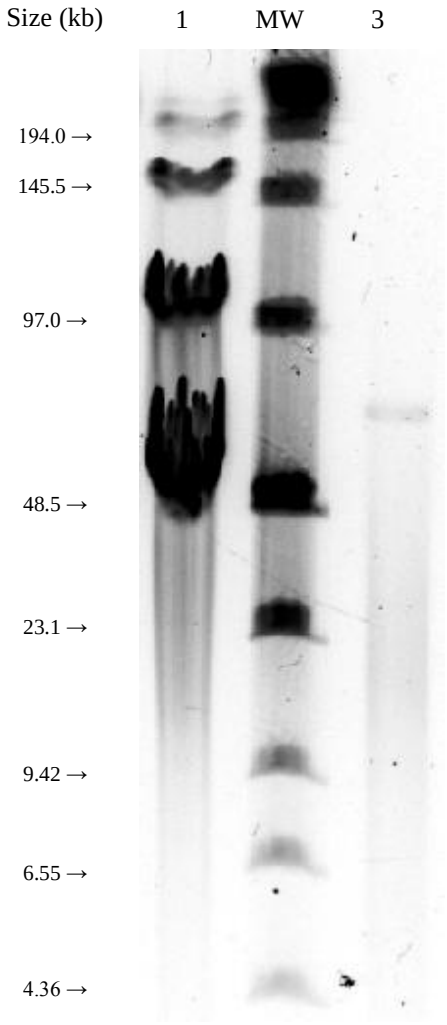
The remaining nine phages belong to the family *Myoviridae*, and both this family and *Siphoviridae* belong to the order Caudovirales. Hatfull has also noted the absence of the family *Podoviridae*, which can be explained either as an evolutionary constrains or a physical constrain in tackling the complex and thick mycobacterial cell wall during infection.

### 3.4 Phage Genome Sizes

Genomic DNA of all 8 phages were extracted from their purified particles and genomes sizes calculated from gels obtained from pulse field electrophoresis.




A




B

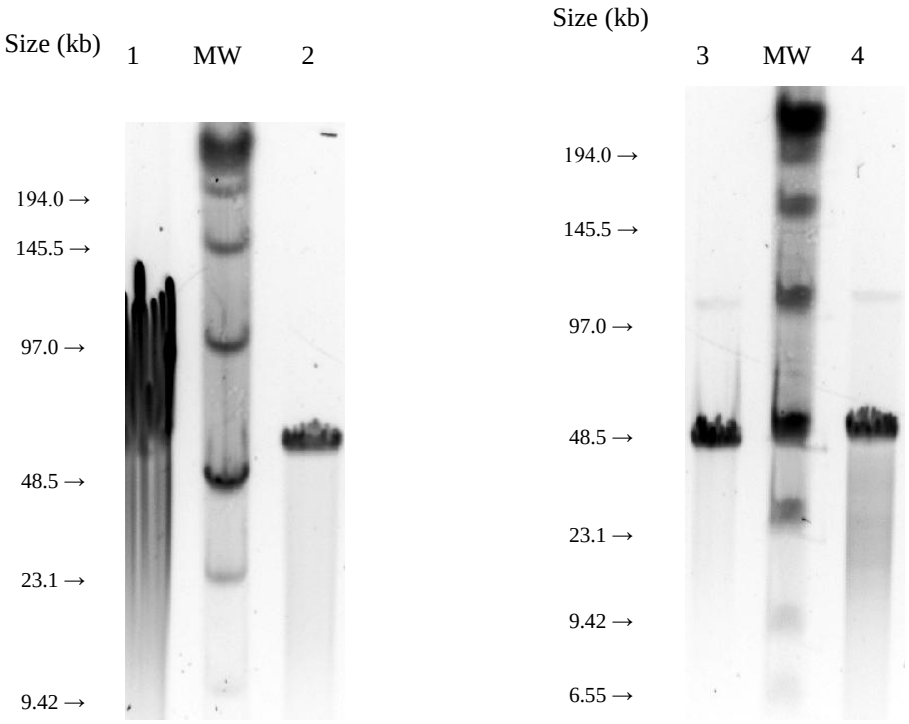



**Figure 3.3** Phage genomic DNA in 1% agarose gel in pulse field gel electrophoresis

A Lane 1: Mozambique 1
Lane 2: Mozambique 3

B Lane 1: Botswana 1A
Lane 3: Botswana 2A

C Lane 1: Caltech 1
Lane 3: Caltech 5

D Lane 1: Botswana 1B
Lane 3: Botswana 2B

The genomic sizes of all 8 phages are given below:

**Table 3.9** Genomic sizes of phages

| **Phage genomic size (kb)** | **Mozambique 1** | **Mozambique 3** | **Caltech 1** | **Caltech 5** |
|---|---|---|---|---|
| | 41 | 39 | 57 | 57 |
| | **Botswana 1A** | **Botswana 1B** | **Botswana 2A** | **Botswana 2B** |
| | 49 | 46 | 67 | 49 |

In a recent review (Hatfull 2010) of 70 sequenced mycobacteriophages, the genome size ranges from 164.60 kb (Myrna) to 41.44 kb (Angel) with the average genome size of 73.60kb with an average of 114.9 ORFs. With the exception of Mozambique 1 and 3, all other isolated phages fall within the range of the 70 genome sizes reviewed. Various studies done on mycobacteriophages have shown that these phages exhibit a diverse and complex set of genomes and it remains to be seen whether there exist mycobacteriophages that has similar genome size to the two outliers found in this study. Botswana 1A genome shows at least four DNA bands of different sizes is due to whole phage genomes binding to each other. The results of Botswana 1A show that these various bands are proportionate relative to each other (See their position relative to marker). Genome sizes are determined by first measuring the distance a given band has travelled and plotted on a log graph of kilo base/base pairs versus distance travelled.

### 3.5 Transduction

Generalized transductions were performed using *M. smegmatis* double mutants that are resistant to both rifampicin and streptomycin. This is because the genes *rpoB* (at position 1461630 – 1465139) and *rpsL* (at position 1498302 – 1498676) are closely linked (Figure 3.3). Rifampicin mutants were produced by plating 200µl of stationary phase *M. smegmatis* $MC^2155$ onto LA media supplemented with 100µg/ml rifampicin and incubated at $37^oC$ for 48hrs or until colonies appear. A single colony was then picked and grown in LB supplemented with 100µg/ml rifampicin and incubated at $37^oC$ for 48hrs or until it reaches stationary phase. 200µl of this liquid culture was in turn plated on a LA plate supplemented with 100µg/ml rifampicin and streptomycin and incubated at $37^oC$ for 48hrs or until colonies appear. These mutants were used to produce plate lysates by incubating phages from liquid lysates with double resistant mutants with cfu/ml of $2 \times 10^6$ and overlaid on LCM agar plates, incubated at $37^oC$ for 48hrs and repeated twice. Phages from the plate lysate were collected using the same method describe above and titred.

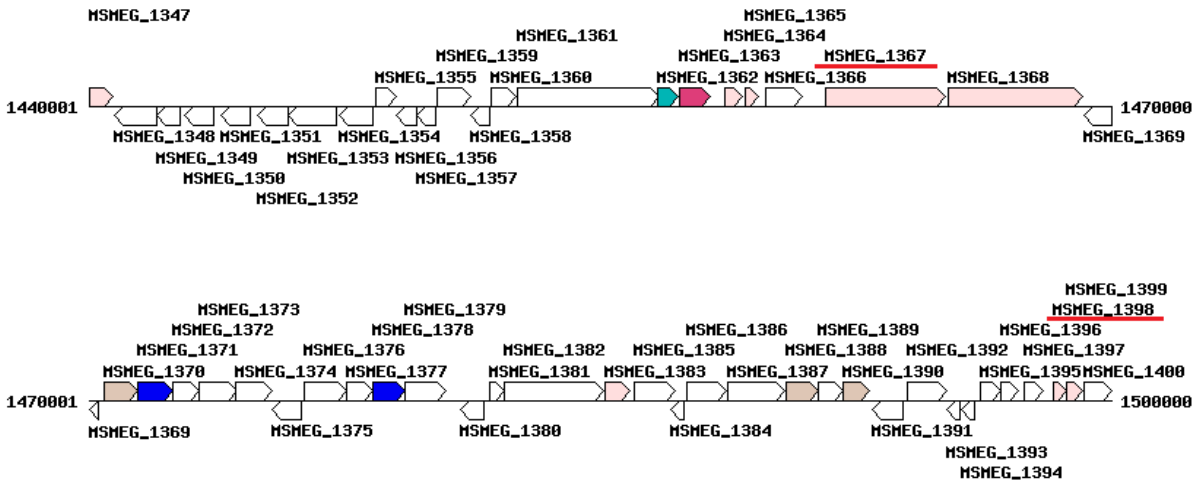



**Figure 3.4** Genome map of *M. smegmatis* showing the locations of *rpoB* (MSMEG_1397) and *rpsL* (MSMEG_1398)

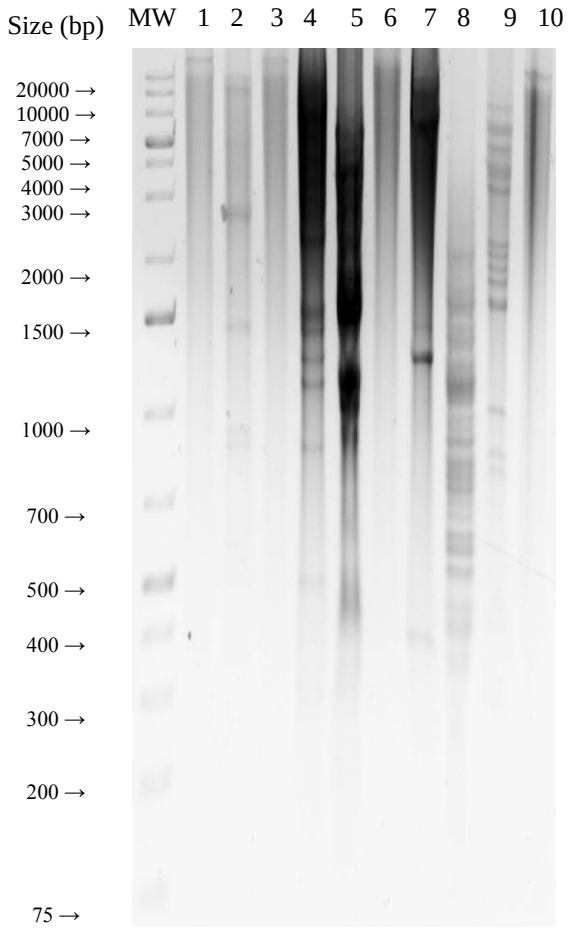
**Table 3.10** Amount phage used in the production of lysate and final titres for transduction

| **Phage** | **Initial titre (pfu) $\times 10^5$** | **MOI** | **Final titre (cfu/ml) $\times 10^9$** |
|---|---|---|---|
| Mozambique 1 | 3 | 0.15 | 4 |
| Mozambique 3 | 5 | 0.25 | 48 |
| Caltech 1 | 20 | 0.01 | 4 |
| Caltech 5 | 1 | 0.05 | 2 |
| Botswana 1A | 60 | 0.03 | 3 |
| Botswana 1B | 1 | 0.05 | 20 |
| Botswana 2A | 70 | 0.35 | 2 |
| Botswana 2B | 1 | 0.05 | 2 |

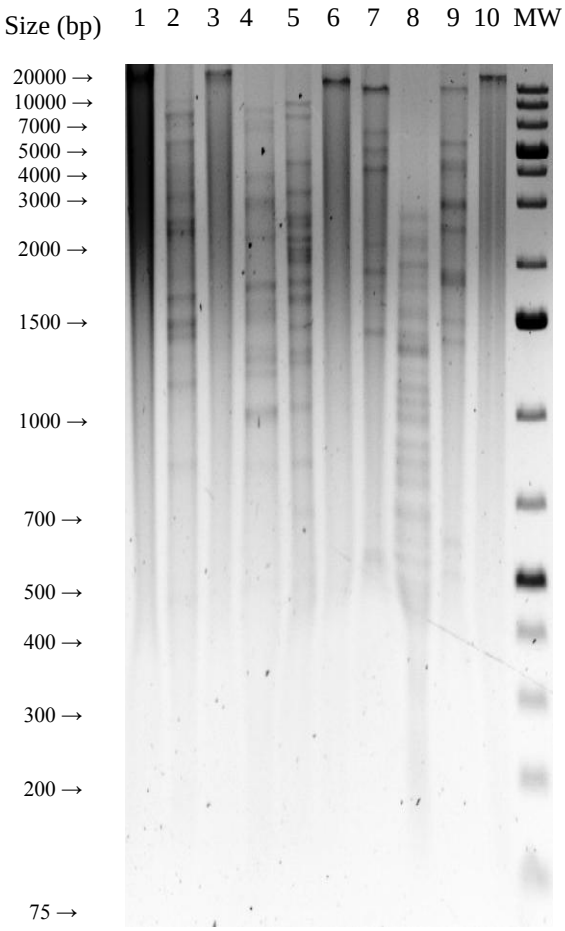
Phages were then mixed with non-resistant *M. smegmatis* at both MOI of 0.1 and 1.0. This mixture was then incubated at room temperature for 30 min. The mixture was then centrifuged for 5 min and the pellet re-suspended and spread on a LCM agar plate containing rifampicin and streptomycin at a concentration of 100µg/ml. Plates were then incubated at 37$^{o}$C for 14 days and observed for bacterial growth every 2 days. No transductants were detected and therefore it could be concluded that the selected phages tested are not generalized transducing phages.

### 3.6 Restriction digest of phage genomic DNA

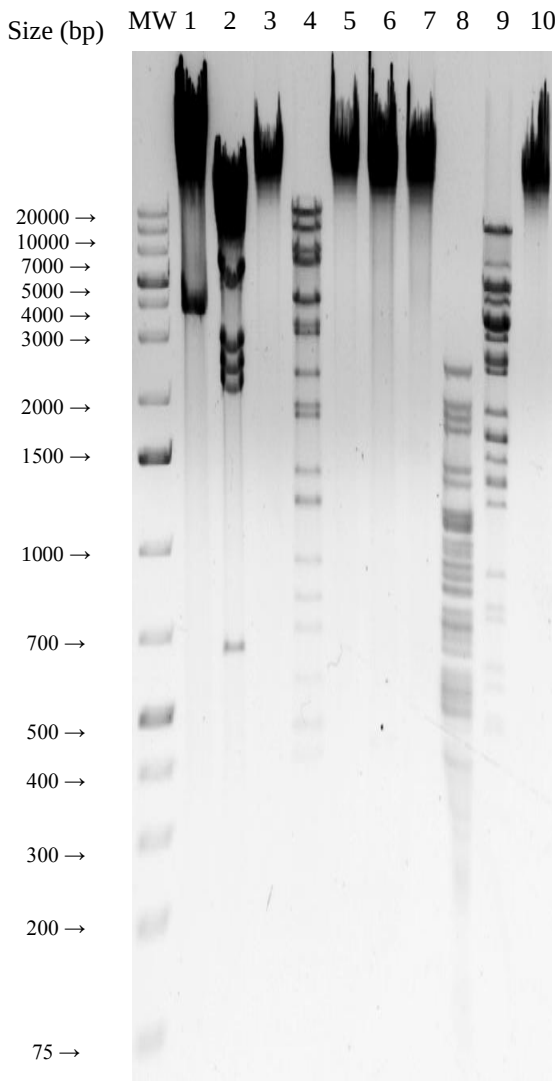
Phage DNA was digested to screen for restriction endonuclease suitable for construction of genomic libraries. Restriction enzymes *Hind*III, *Bgl*II, *Pst*I, and *Sfu*I were tested, also included were isocaudomers of these restriction enzymes, such as *Bcl*I, *Nsi*I, *Bam*HI and *Xmi*I. Results for all digestions are shown below.



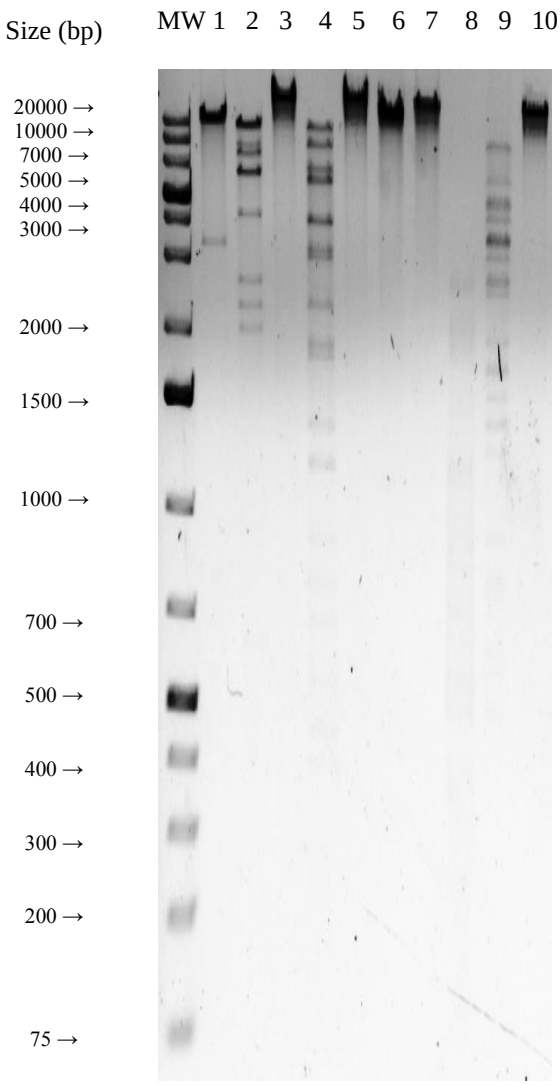

Mozambique 1



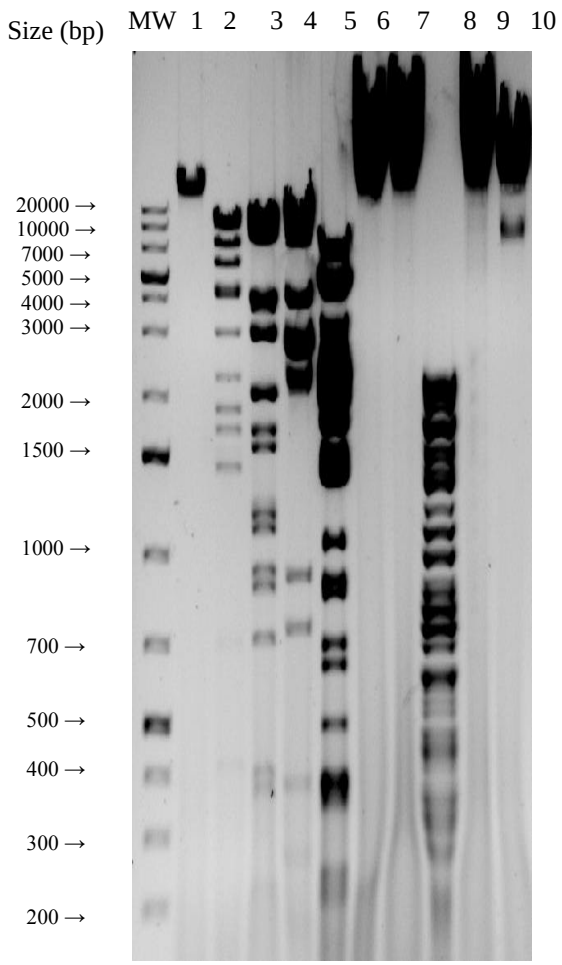

Mozambique 3



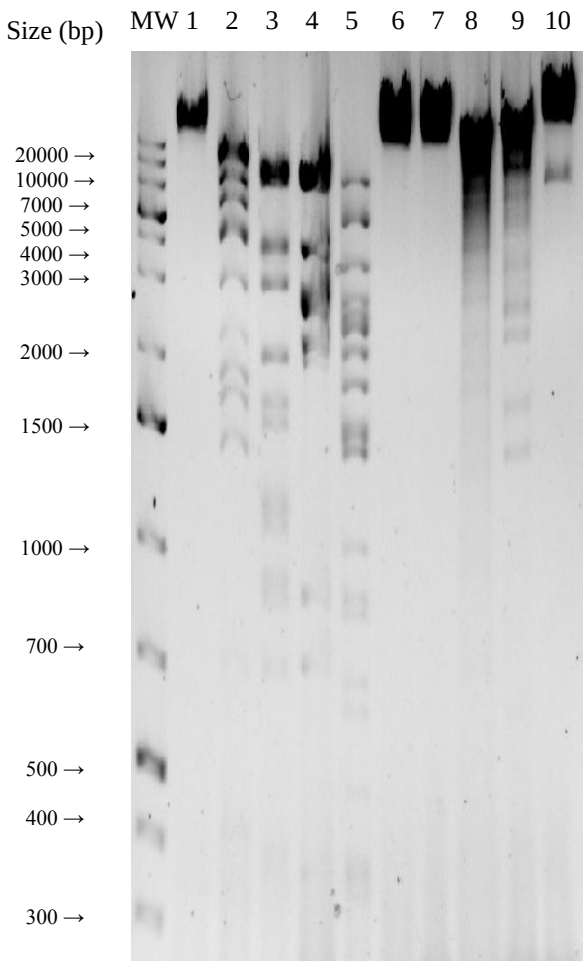

Caltech 1




Caltech 5




Botswana 1A




Botswana 1B

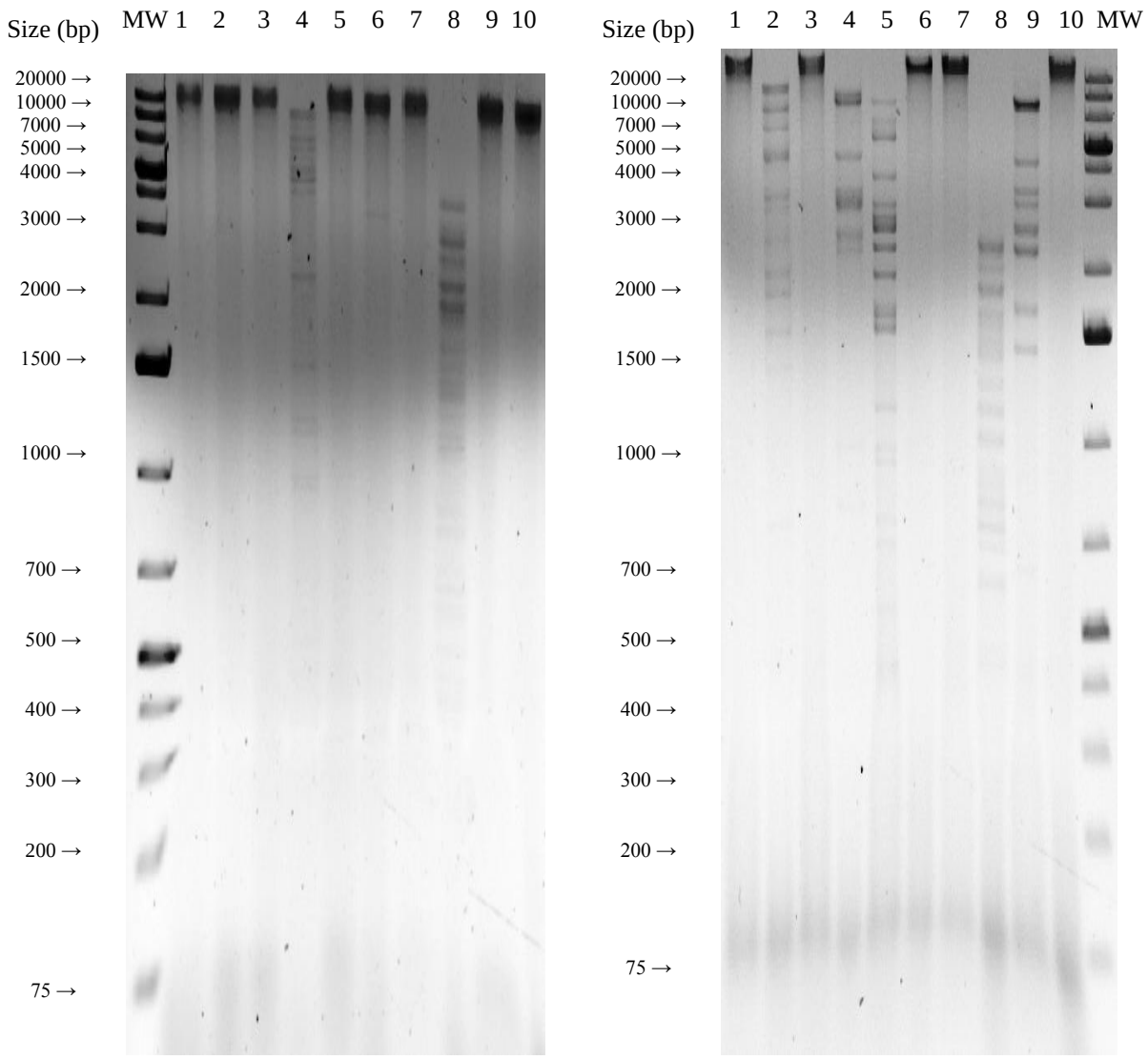



**Figure 3.5** Digestions of phage genomic DNA by various restriction enzymes on 0.8% agarose gel. Lane numbers corresponds to the same enzymes on all gels. 1 *Hind*III; 2B*gl*II; 3 *Bam*HI; 4 *Bcl*I; 5 *Pst*I; 6 *Nsi*I; 7 *Sfu*I; 8 *Xmi*I; 9 *Cla*I; 10 Undigested control

Throughout the co-evolutionary history that bacteria have with phages, host cells have evolved a series of defense mechanisms to resist phage infection. One such mechanism is the restriction-modification, where foreign DNA (such as that from a phage) is degraded by host restriction enzymes (RE) upon entry into the cell. The host own genetic material will be protected by methylation through bacterial methylase. As such, phages have developed

several resistance mechanisms, such as the development of unusual bases, modification of bases and altering its sequence in order to defeat host restriction, with most of the research done on *E.coli* and *B. subtilis* phages

For unusual bases, *Bacillus subtilis* phages PBS1And PBS2 incorporate uracil instead of thymine to protect the genome from degradation (Takahashi and Marmur 1963). The two phages encodes the protein Ugi, which mimics uracil containing double stranded DNA and interacts with host protein uracil-DNA glycosylase, a host protein that counteracts this sequence alteration by removing uracil bases from phage DNA (Wang and Mosbaugh 1989).

Modification of bases is another strategy employed by phages. An example of which can be found in phage T4, which methylates adenosine residues (Carlson *et al* 1994) and protects the genome from classical RE defence mechanisms. In addition, the phage replaces cytosine with a modified hydroxymethylcytosine (HMC) in its genome before being glycosylated. This protects the HMC from degradation by host Mcr restriction enzymes that were developed to cleave modified cytosine (Bickle and Kruger 1993). In response, another RE, GmrSD, was developed to specifically recognize and cleave glycosylated HMC and this lead T4 to encode for an internal head protein, IPI* to protect the genome against GmrSD through direct protein-protein interaction when it is injected into the host along with the genome (Bair *et al* 2007). A recent study indicates that host cell produces GmrSD that are insensitive to IPI* (Rifat *et al* 2008), highlighting a robust defence mechanism in response to phage attacks.

For mycobacteriophages, restriction studies (Coene *et al* 1993) done on a series of *M. smegmatis* hosts shows phages such as 29M, 31M and 122 to have methylcytosine in the sites GGCC and CCGG, showing resistance to enzyme *Hae*III and *Hpa*III. Other phages such as 154 and 110 are resistant to *Sau*3AI, *Mbo*I and *Hha*I due to the presence of methylated cytosine and adenine

### 3.7 Construction of Phage Genomic Libraries

Four genomic libraries were constructed using three of the phages: Mozambique 1, Mozambique 3, and Botswana 2A. All libraries were constructed by transforming *E. coli* MM294-4 with pDA71And 20 clones from each library were randomly chosen (with the exception of Mozambique 3 *Pst*I library, from which 60 clones were randomly chosen).

#### 3.7.1 Mozambique 1 *Pst*I Library

Genomic library of phage Mozambique 1 was constructed with endonuclease *Pst*I, and 20 clones were randomly chosen and plasmid mini preparation was done to extract these clones.

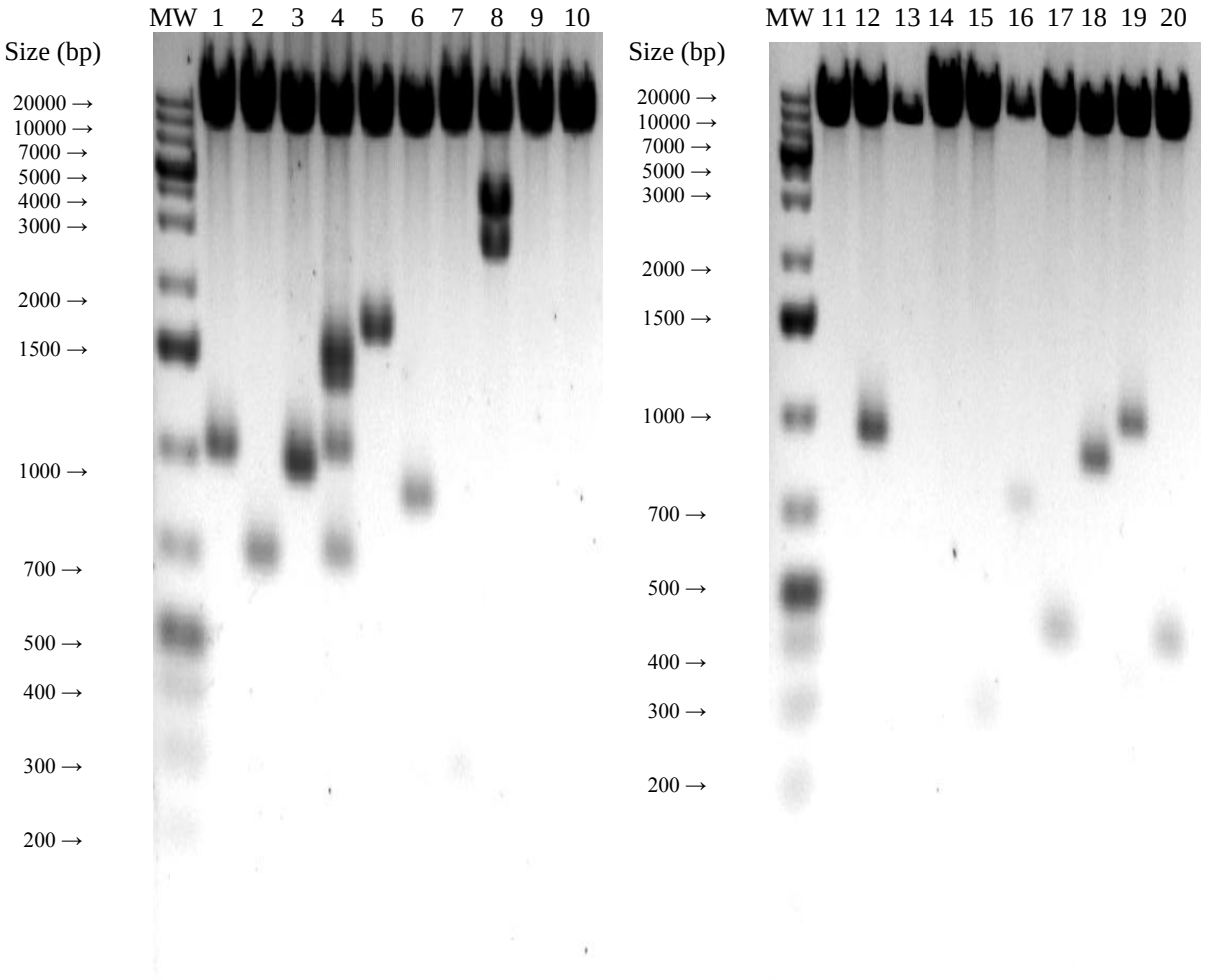


**Figure 3.6** Twenty clones from Mozambique 1 *Pst*I library shown on two 0.8% agarose gels. Inserts are bands below the upper row of bands, which are linearised pDA 71.

Once the clones are obtained, calculations were performed to determine the average insert size and number of clones needed to have 95% probability of the library containing the Mozambique 1 genome.

Average insert size = Sum of all insert size/number of clones isolated

= 22470 bp/20

= 1.124 kb

To calculate the number of clones needed to have 95% probability of the library containing the phage genome, the following equation is used

$$N = \frac{\ln (1 - p)}{\ln (1 - a/b)}$$

$$= \frac{\ln (1 - 0.95)}{\ln (1 - 1.124/40.59)}$$

= 105 clones

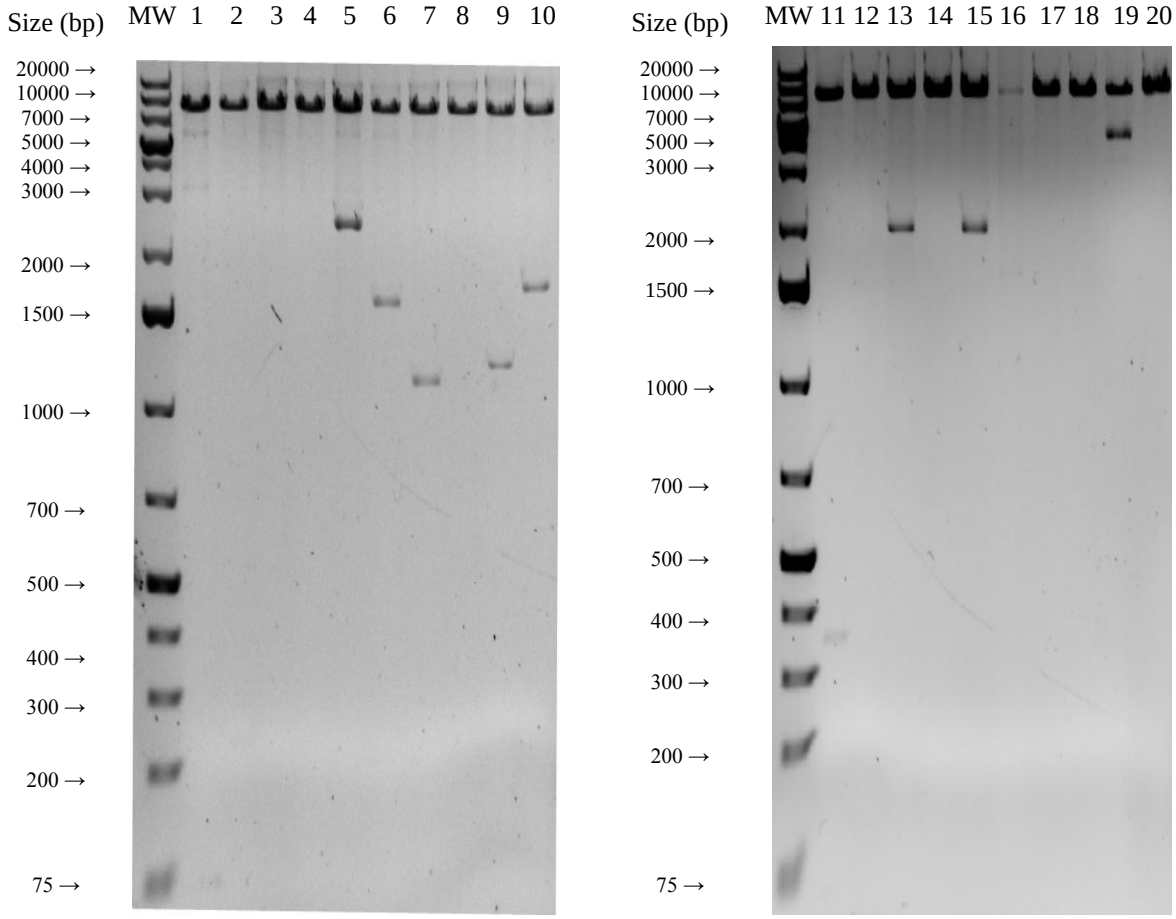
The number of clones obtained from this library was 141.

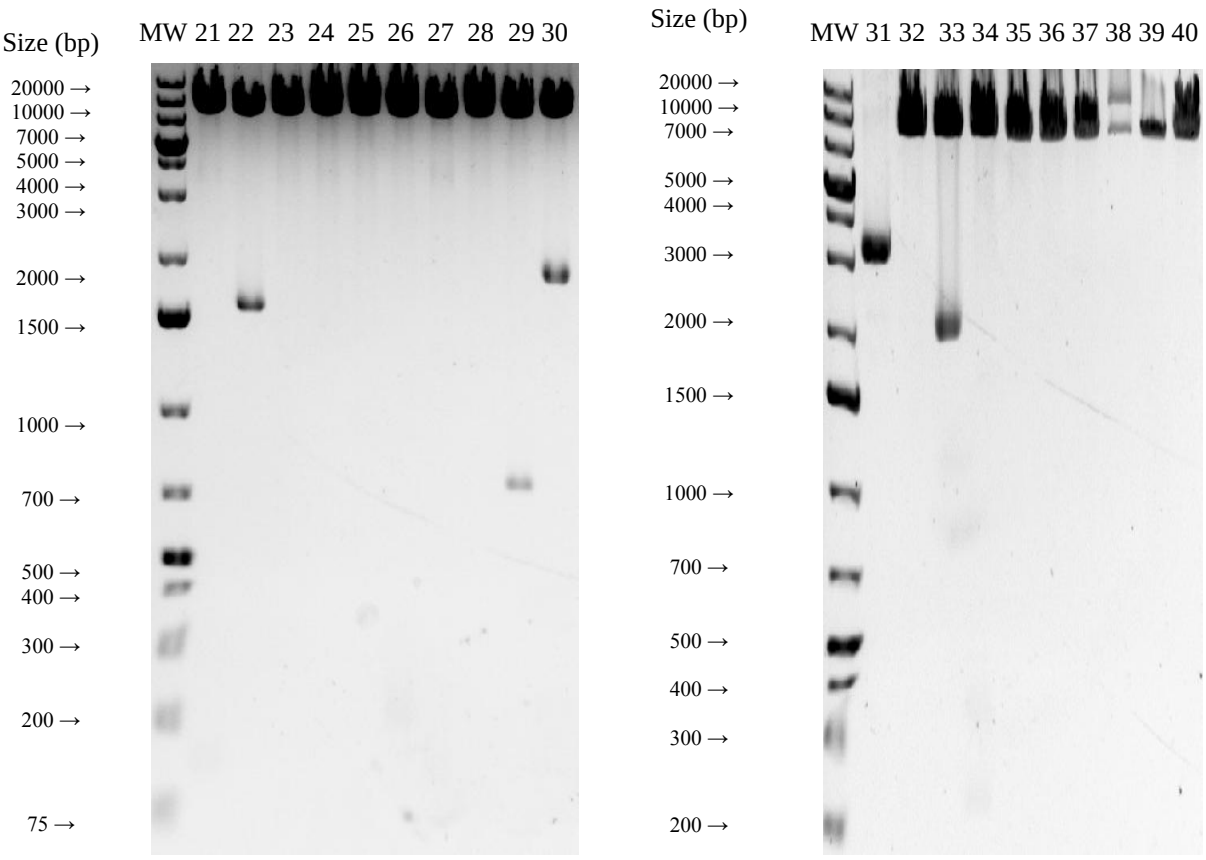
Where N is the number of clones, p is the probability of any phage DNA in the library and a and b are average insert size and the total phage genome size respectively.

### 3.7.2 Mozambique 3 *Pst*I Library

Genomic library of phage Mozambique 3 was constructed with endonuclease *Pst*I.

Sixty clones were randomly chosen and plasmid mini preparation was done to extract these clones




Size (bp)
MW 21 22 23 24 25 26 27 28 29 30
20000 →
10000 →
7000 →
5000 →
4000 →
3000 →
2000 →
1500 →
1000 →
700 →
500 →
400 →
300 →
200 →
75 →
Size (bp)
MW 31 32 33 34 35 36 37 38 39 40
20000 →
10000 →
7000 →
5000 →
4000 →
3000 →
2000 →
1500 →
1000 →
700 →
500 →
400 →
300 →
200 →

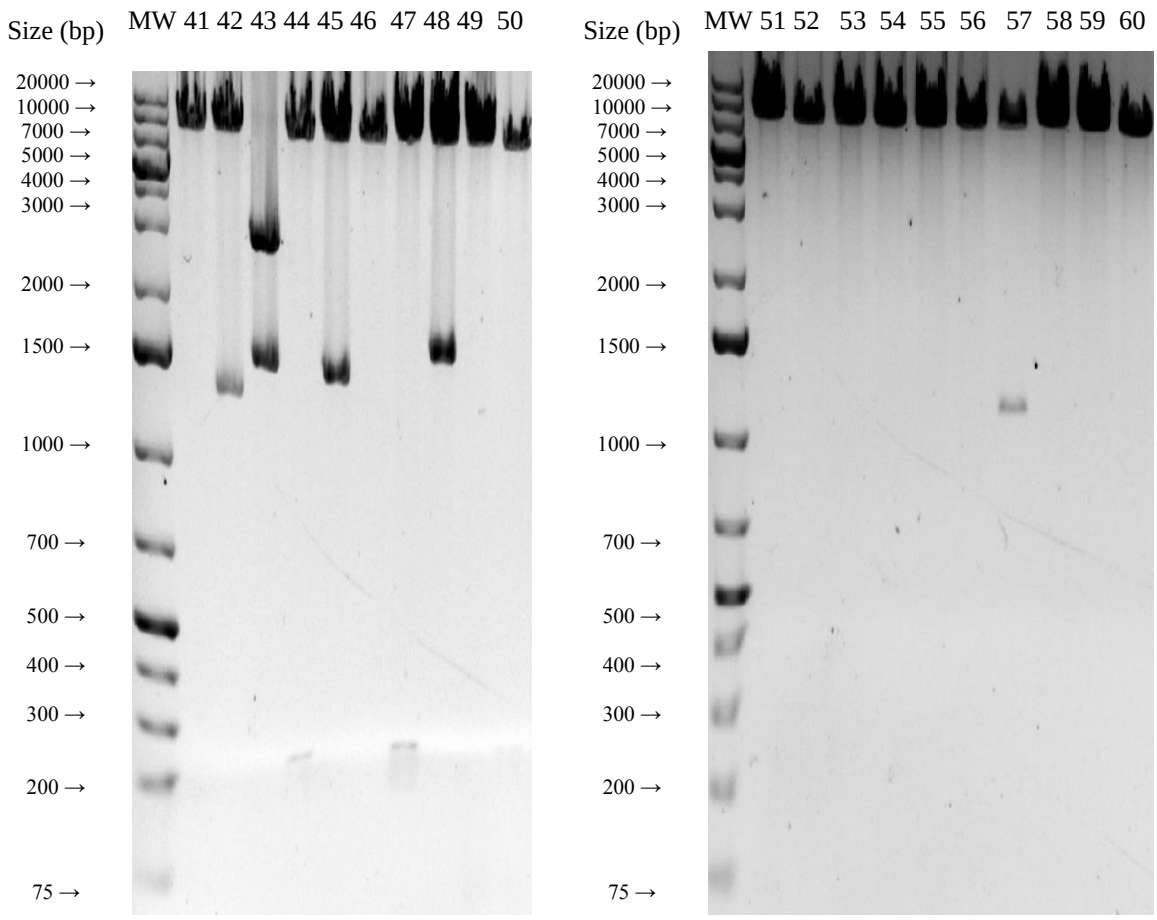



**Figure 3.7** Sixty clones from Mozambique 3 *Pst*I Library shown on two 0.8% agarose gels. Inserts are bands below the upper row of bands, which are linearised PDA 71.

The average insert size was 0.47kb with 247 transformants needed to have a 95% probability of the library containing the Mozambique 3 genome. A total of 736 transformants were obtained from this library.

### 3.7.3 Mozambique 3 *Bam*HI Library

Genomic library of phage Mozambique 3 was constructed with endonuclease *Bam*HI.

Twenty clones were randomly chosen and plasmid mini preparation were done to extract these clones

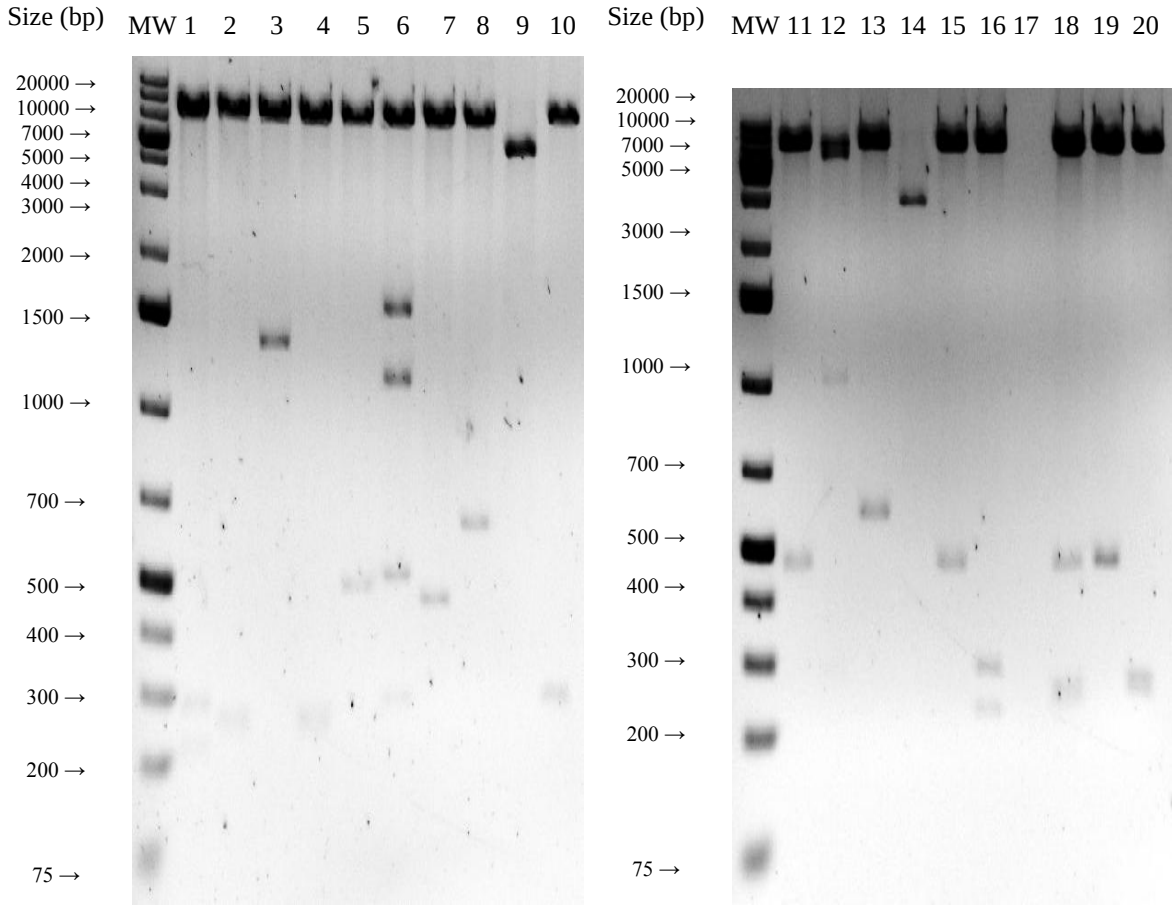



**Figure 3.8** Twenty clones from Mozambique 3 *Bam*HI Library shown on two 0.8% agarose gels. Inserts are bands below the upper row of bands, which are linearised vector PDA 71

The average insert size was 0.24kb with 499 transformants needed to have a 95% probability of the library containing the Mozambique 3 genome. A total of 601 transformants were obtained form this library.

### 3.7.4 Botswana 2A *Bcl*I Library

Genomic library of phage Botswana 1A was constructed with endonuclease *Bcl*I.

Twenty clones were randomly chosen and plasmid mini preparation were done to extract these clones

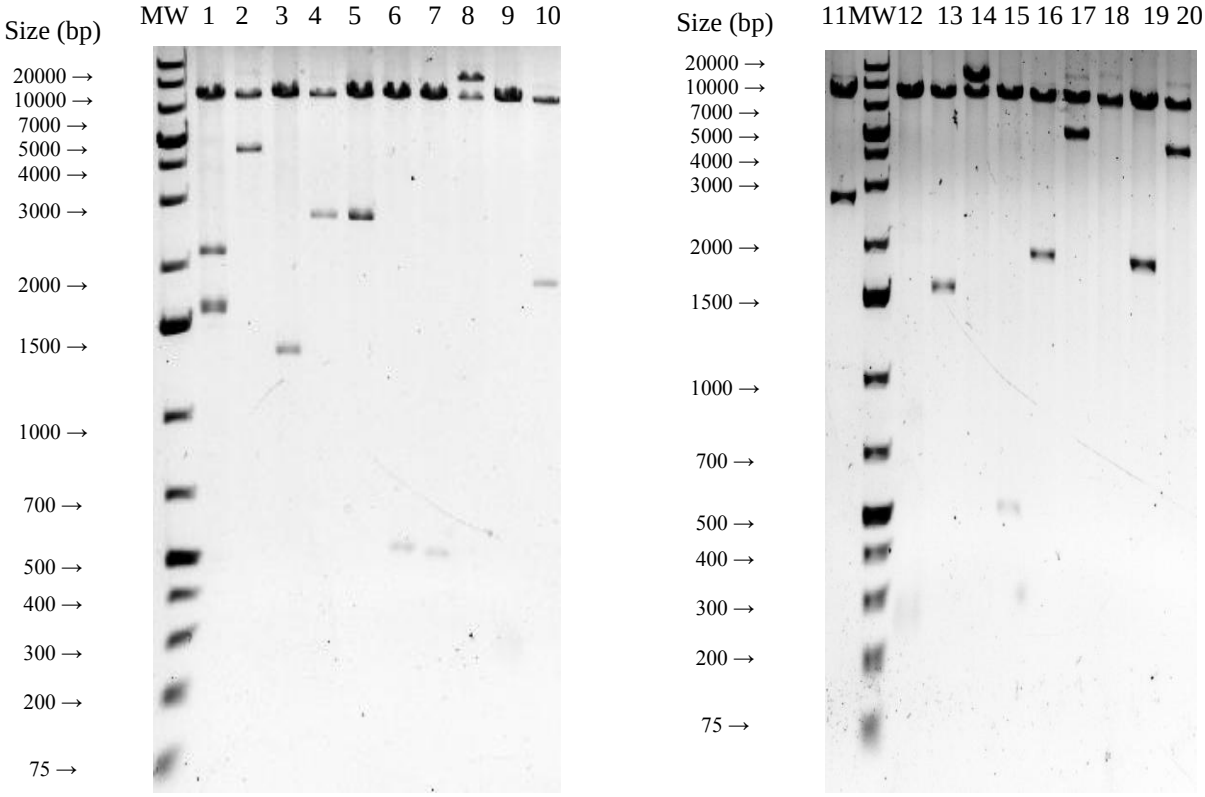


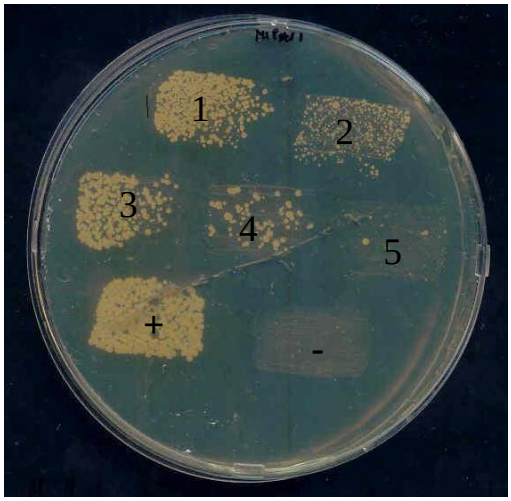

**Figure 3.9** Twenty clones from Botswana 2A *Bcl*I Library shown on two 0.8% agarose gels. Inserts are bands below the upper row of bands, which are linearised vector
PDA 71 with the exception of clone nos. 8 and 14, in which case the insert is positioned above the linearised vector.

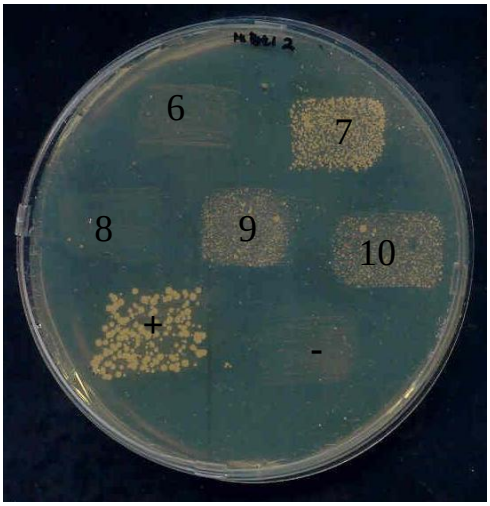
The average insert size is 2.924kb with 67 transformants needed to have a 95% probability of the library containing the Botswana 2A genome. A total of 642 transformants were obtained for this library.

### 3.8 Screening of Libraries for Clones with Antimicrobial Properties

Clones from libraries made from phages Mozambique 1And 3, along with phage Botswana 1A were screened for antimicrobial activity by transforming these clones into *R. erythropolis* SQ1By PEG mediated transformation. With the libraries of phage Mozambique 1And Botswana 1A, all twenty clones were screened. With Mozambique 3, 11 clones with inserts were selected screened. The modified vector pDA71*, which is a pDA71 vector with the *EcoR*I suicide gene removed, was used as the positive control.

#### 3.8.1 Screening of Mozambique 1 *Pst*I Library Clones with Antimicrobial Properties



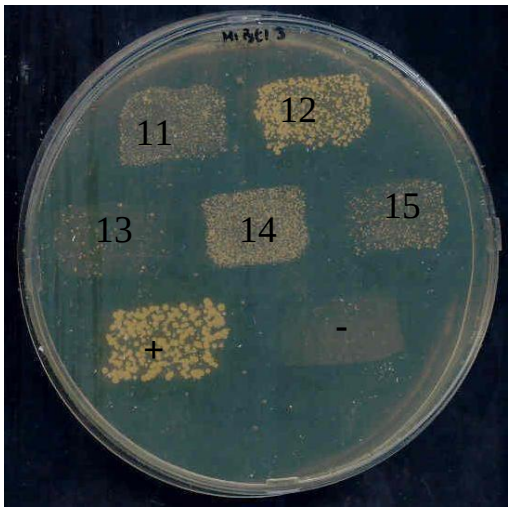




**Fig 3.10** PEG mediated transformation of *R. erythropolis* SQ1 of 20 clones isolated from *Pst*I Mozambique 1 library. Numbers on the plate corresponds to the numbers of the clone. "+" denotes the positive control of pDA71* and "-"denotes negative control

**Table 3.11** Transformation efficiencies for twenty clones isolated from Mozambique 1 *Pst*I library

| Clone # | Transformants/µg of DNA |
|---|---|
| 1 | 441 |
| 2 | 76 |
| 3 | 156 |
| 4 | 32 |
| 5 | 3 |
| pDA71* control | 2027 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 6 | < 1.64 |
| 7 | 385 |
| 8 | < 0.9 |
| 9 | 2 |
| 10 | 16 |
| pDA71* control | 940 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 11 | 11 |
| 12 | 822 |
| 13 | 104 |
| 14 | 750 |
| 15 | 387 |
| pDA71* control | 1000 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 16 | 7 |
| 17 | 41 |
| 18 | 7 |
| 19 | 125 |
| 20 | 616 |
| pDA71* control | 767 |

### 3.8.2 Screening of Selected Clones from Mozambique 3 *Pst*I Library for Antimicrobial Properties

Of the clones isolated from phage Mozambique 3 *Pst*I library, only selected clones with inserts were screened for antimicrobial activities.

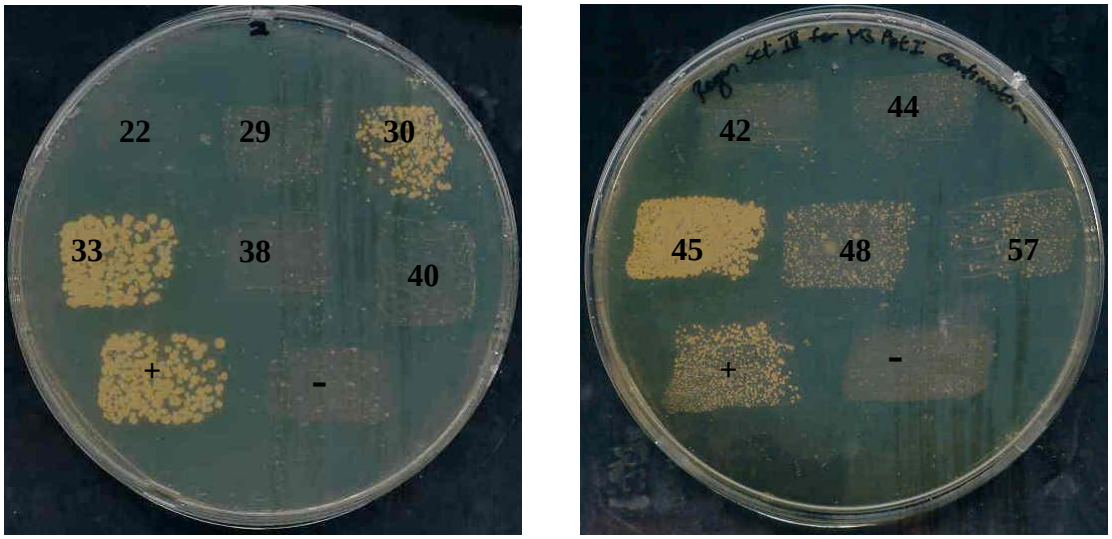



**Figure 3.11** PEG mediated transformation of *R. erythropolis* SQ1 of 11 selected clones isolated from *Pst*I Mozambique 3 library. Numbers on the plate corresponds to the numbers of the clone. "+" denotes the positive control of pDA71* and "-"denotes negative control

**Table 3.12** Transformation efficiencies for selected clones isolated from Mozambique 3 *Pst*I library

| Clone # | Transformants/µg of DNA |
|---|---|
| 22 | < 3 |
| 29 | < 4 |
| 30 | 743 |
| 33 | 1522 |
| 38 | 20 |
| 40 | < 3 |
| pDA71* Control | 860 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 42 | < 12.5 |
| 44 | 5 |
| 45 | 1174 |
| 48 | 748 |
| 57 | 18 |
| pDA71* control | 1447 |

### 3.8.3 Screening of Botswana 2A *Bcl*I Library Clones with Antimicrobial Properties

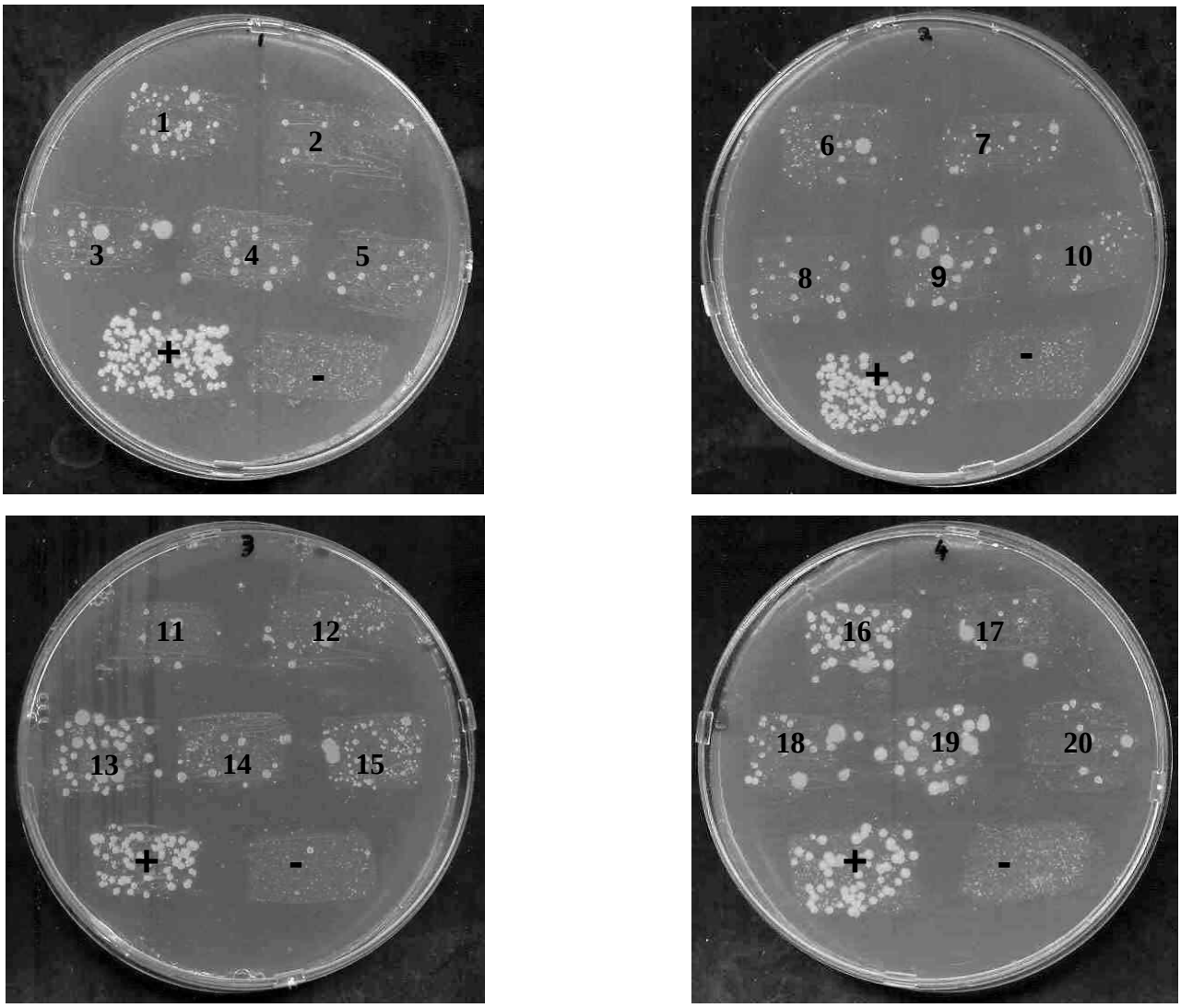



**Figure 3.12** PEG mediated transformation of *R. erythropolis* SQ1 of 20 clones isolated from *Bcl*I Botswana 2A library. Numbers on the plate corresponds to the numbers of the clone. "+" denotes the positive control of pDA71* and "-"denotes negative control.

**Table 3.13** Transformation efficiencies for 20 clones isolated from Botswana 1A *Bcl*I library

| Clone # | Transformants/µg of DNA |
|---|---|
| 1 | 291 |
| 2 | 67 |
| 3 | 110 |
| 4 | 205 |
| 5 | 90 |
| pDA71* control | 873 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 6 | 110 |
| 7 | 80 |
| 8 | 123 |
| 9 | 188 |
| 10 | 71 |
| pDA71* control | 500 |

| Clone # | Transformants/µg of DNA |
|---|---|
| 11 | 89 |
| 12 | 113 |
| 13 | 449 |
| 14 | 60 |
| 15 | 514 |
| pDA71* control | 413 |

| Clone # | Transformants/μg of DNA |
|---|---|
| 16 | 400 |
| 17 | 72 |
| 18 | 314 |
| 19 | 313 |
| 20 | 100 |
| pDA71* control | 593 |

### 3.9 Analysis of Inhibitory Clones

Several clones that showed inhibitory activities from *Pst*I libraries of phages Mozambique 1And 3 were selected for further analysis. Inserts from these clones were extracted from 0.6% agarose gel, purified, ligated into vector pUC18/19. These inserts in pUC18/19 was then transformed into *E. coli* MM294-4 and sequenced. Sequencing was done by Inqaba Biotech.

Sequencing was done from both forward and reverse direction to ensure the coverage of the entire insert. Forward and reverse sequences was aligned through EMBOSS Needle global pairwise alignment. The sequence was then analyzed using BLASTn to determine the nucleotide sequence similarity of each clone to those on the database and BLASTx was used to determine the amino acid sequence similarity.

Open Reading Frames (ORFs) in each clone was then predicted using GeneMark and Softberry. The former is a heuristic approach using small amount of prokaryotic nucleotide sequence to build up various Markov models. This can be used to find ORFs in small fragments of sequences from unknown prokaryotic organisms to viruses and phages and even organelles and plasmids (Yok and Rosen 2011). The latter also uses a Markov chain-based gene prediction in viruses. The ORFs reported below are the results taken from both programs.

Once the ORFs are predicted, protein sequences of the ORFs were analyzed using BLASTp to determine protein-protein sequence similarities with those in the database. Clones were also analyzed for restriction sites using NEB cutter in order to determine the minimum amount of DNA needed for inhibitory activities. This was done to screen for any relevant enzymes that can be used to digest the inserts in order for these fragments to be re-cloned back into vector pDA71 for transformation into *R. erythropolis* SQ1. Restriction endonuclease screened is the same ones used in various digest of phage DNA for the construction of phage genomic libraries.

### 3.9.1 Clone 5 of Phage Mozambique 1 *Pst*I Library

Insert of clone 5 were extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below

ctgcagtggc gccgcgtarg acgggttgca ccagacgcgg atgtgcttgg tctgagccat gtcgtgcaaa aacmcgmcct cgaggtaagc gtcaccggaa tcggtcttga cgacccggta gtggtccatc atcccggtcc accgagcgcc ttgcttctcg atgttgatga tgacgttccg ctttgctcga ccacggtggt tcatgaccca cttggccagg tagtggctca gcgagagctg aagggtcgcg gtgcccacgt cgttctcgat gaactcccac tcgagcaacc gctcgccagc gaccacgcct cggaggcgga agtcgccgtc gcgcagctct acgtcggccg gtgccagacg tgcggcctcg cgcttggcgc ggcgctgttg gatgagcttc caragatcgt tggcctgaga gactgacttc agaccactca ctcgagcccc cagcagcgcg tccacggcct gggaagccga agggtcacta cctgtcccgg agcgcatccc gaggcgtcta tgacgaactc agcctcttcg gtgtacggcg ggatctggtt gcggaaccgg acgccgttca tgcgagccca cacaggcgag cccgactcgg aagcgatctg ctcctcgcgg cgatcggtgt cgatcacgca gttctcgccg tagatcaagc cagggagctt gagccggcgg ttggcggaaa ctccttcgtt cttcgatcga agtaagttca ggaatgacga actgagtgaa tcggagctgt ttcccacggg atcggcgtgt tcggagggaa cggcccaggr ggaagttggg caccttctcg gtggagccgg gaacggtcca cttcggggcg atgtactggt cggtggggtt gagcccacct tgtccacggc cgaccttgat ccgcagcgtc tccttgggca gttcctccca cggccacgg cggcgtccag aacgacgggt cgaaccgggt gtcggtcttc gtcttggccg agaagacgcg atcgtycttc ccaccagaac gggtcgtacg cgatgcacga catccaccgg tcagggttga tgctgtttgc ctcgcgggtc ggtctttcat cytcgacttg gggggactcg aacaggaccg gcaccttgaa ggtagcgcgt accggagtcc ggggtggtg acgtagatct tgcagtcgcg gttgaacgcc cacgccttgc gccactcgc tgtcacgcga gagccacgac cggggtccgc tcttggcatc gttgaggatc tgaaccccga agacgatgtc ccgcttcagg atccggtgcg acaagtagcg agcgccgggg tagttccccg gctcctcaat cacgaccttg acgggagggt cgtagaaaca accctccacg tctgtggcca ggaacacgcc ctggtcaccg gtcgtcagat tgaagaactc accattgaca ccctcgagtt caacgatggt gtcggtgatc aatgcttacc tcctggtgaa tgtcaagttc gcggcctact gacggccgac gactgcgagt gcgctcttgc tctcttcgcg gtccttgatc gacagagcct catcgaccga gccgatgttg aacacgtacg agatgccctc ggtgatggcc ttcgagatga atccctcacc agagatgccg atgtccgaaag gaactgcttg ccagtcgctt tcgcgaagtc gaccggggag gccatcatct tggcgacttg gtcctccagc gacataccct tcgatccctc agcggtctcg tcgkkgtagt cgcctacgag gtcgagcaga tccttttgag cctgcag

**Figure 3.13** Sequence from clone 5 of Mozambique 1 *Pst*I library with *Pst*I site highlighted in orange.

BLASTn result

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| JN408461.1 | Mycobacterium phage Trixie, complete genome | 1810 | 1810 | 98% | 0.0 | 85% |
| Z18946.1 | Mycobacteriophage L5 complete genome | 1704 | 1704 | 92% | 0.0 | 85% |
| JN408460.1 | Mycobacterium phage Turbido, complete genome | 1668 | 1668 | 98% | 0.0 | 84% |
| GU339467.1 | Mycobacteriophage RedRock, complete genome | 1629 | 1629 | 98% | 0.0 | 84% |
| AF022214.1 | Mycobacteriophage D29, complete genome | 1505 | 1505 | 92% | 0.0 | 83% |
| DQ398043.1 | Mycobacteriophage Che12, complete genome | 1442 | 1442 | 93% | 0.0 | 82% |
| EU744250.1 | Mycobacterium phage Pukovnik, complete genome | 1322 | 1322 | 93% | 0.0 | 81% |
| JF937094.1 | Mycobacterium phage Hammer, complete genome | 1185 | 1185 | 94% | 0.0 | 80% |
| JF704097.1 | Mycobacterium phage Gladiator, complete sequence | 1164 | 1164 | 94% | 0.0 | 80% |
| JF937092.1 | Mycobacterium phage DaVinci, complete genome | 1094 | 1094 | 94% | 0.0 | 79% |
| JN049605.1 | Mycobacterium phage EricB, complete genome | 1094 | 1094 | 94% | 0.0 | 79% |
| JF704110.1 | Mycobacterium phage PackMan, complete sequence | 888 | 1161 | 90% | 0.0 | 81% |

**Figure 3.14** BLASTn search result from clone 5 of Mozambique 1 *Pst*I library when aligned with known sequences from the database. Where "scores" are similarities between the query and the matches found in the database and "Total score" is the sum of all scores between the query and high scoring hits, "Max score" is similar to E (Expected) value. This is the number of hits one is "expected" to see when searching the database for a particular length that is the same as the query. This is dependent on both the query length and hits in the database. The closer the E-value is to zero, the more significant the hit. "Coverage" deals with the percentage of coverage of the query and "Max Ident" is the highest percentage identity of highly similar pairs.

BLASTx result

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| AEL17860.1 | gp30 [Mycobacterium phage Trixie] | 238 | 410 | 36% | 7e-71 | 92% |
| AEL17766.1 | gp29 [Mycobacterium phage Turbido] | 234 | 453 | 41% | 2e-69 | 97% |
| NP_039691.1 | observed 38.5Kd protein [Mycobacterium phage L5] >sp\|Q05234.1\|V | 231 | 449 | 41% | 3e-68 | 95% |
| AEK07301.1 | gp28 [Mycobacterium phage PackMan] | 231 | 437 | 41% | 4e-68 | 90% |
| NP_046843.1 | gp27 [Mycobacterium phage D29] >sp\|O64221.1\|VG27_BPMD2 RecNa | 229 | 446 | 41% | 2e-67 | 92% |
| YP_001994846.1 | gp29 [Mycobacterium phage Pukovnik] >gb\|ACE79955.1\| gp29 [Mycc | 229 | 437 | 41% | 2e-67 | 89% |
| YP_655608.1 | gp29 [Mycobacterium phage Che12] >gb\|ABE67348.1\| gp29 [Mycoba | 228 | 441 | 41% | 4e-67 | 96% |

**Figure 3.15** BLASTx search result from clone 5 of Mozambique 1 *Pst*I library when aligned with known protein sequences from the database.

ORF prediction and construction of restriction map

Open reading frames were predicted using both GeneMark and SoftBerry, a given ORF is accepted when there is a consensus between the results from the two programs. Clone 5 has two ORFs, with ORF1 occupying the position between 1086 to1469 and ORF2 from 1503 to 1805. A restriction map of the insert is also drawn after analyzing the sequence using NEB cutter showing the location of the restriction sites and the positions of the two ORFs.

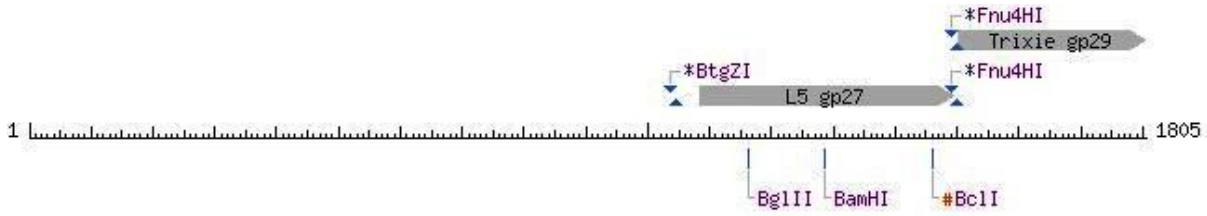



**Figure 3.16** Restriction map of clone 5 of Mozambique 1 *Pst*I library along with the positions of the two predicted ORFs

BLASTp results of clone 5 ORFs

The two ORFs of clone 5 were analyzed to determine any similarity to proteins in the database; both L5 gp27 and Trixie gp29 are minor tail subunits.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| NP_039691.1 | observed 38.5Kd protein [Mycobacterium phage L5] >sp\|Q05234.1\|V | 121 | 121 | 100% | 7e-33 |
| AEK07301.1 | gp28 [Mycobacterium phage PackMan] | 120 | 120 | 100% | 2e-32 |
| AEL17860.1 | gp30 [Mycobacterium phage Trixie] | 120 | 120 | 100% | 2e-32 |
| AEL17766.1 | gp29 [Mycobacterium phage Turbido] | 119 | 119 | 100% | 3e-32 |
| YP_001994846.1 | gp29 [Mycobacterium phage Pukovnik] >gb\|ACE79955.1\| gp29 [Mycc | 119 | 119 | 100% | 3e-32 |
| YP_655608.1 | gp29 [Mycobacterium phage Che12] >gb\|ABE67348.1\| gp29 [Mycoba | 119 | 119 | 100% | 5e-32 |
| NP_046843.1 | gp27 [Mycobacterium phage D29] >sp\|O64221.1\|VG27_BPMD2 RecNa | 118 | 118 | 100% | 1e-31 |
| ADB93723.1 | gp30 [Mycobacterium phage RedRock] | 117 | 117 | 100% | 2e-31 |
| AEK08676.1 | gp30 [Mycobacterium phage Hammer] | 115 | 115 | 100% | 8e-31 |
| AEJ95031.1 | gp30 [Mycobacterium phage Gladiator] | 115 | 115 | 100% | 8e-31 |
| AEJ93325.1 | gp30 [Mycobacterium phage EricB] >gb\|AEK08473.1\| gp30 [Mycobac | 115 | 115 | 100% | 1e-30 |

ORF1

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEL17859.1 | gp29 [Mycobacterium phage Trixie] | 161 | 161 | 100% | 6e-45 |
| AEL17765.1 | gp28 [Mycobacterium phage Turbido] | 160 | 160 | 99% | 1e-44 |
| ADB93722.1 | gp29 [Mycobacterium phage RedRock] | 150 | 150 | 100% | 3e-41 |
| NP_039690.1 | predicted 86.4kd protein; 52Kd observed [Mycobacterium phage L5] | 128 | 128 | 99% | 2e-33 |
| YP_655607.1 | gp28 [Mycobacterium phage Che12] >gb\|ABE67347.1\| gp28 [Mycoba | 125 | 125 | 99% | 2e-32 |
| YP_001994845.1 | gp28 [Mycobacterium phage Pukovnik] >gb\|ACE79954.1\| gp28 [Mycc | 123 | 123 | 99% | 1e-31 |

ORF2

**Figure 3.17** BLASTp search result of ORFs 1 and 2 from clone 5 of Mozambique 1 *Pst*I library when aligned with known protein sequences from the database. Blast results for all BLASTp (protein-protein) query does not provide "Max Ident".

Tail like bacteriocins produced by both gram positive and negative bacteria, have been previously described in several studies (Gratia 1987; Krügel *et al* 1987; and Muňoz *et al* 1987). Two types of such bacteriocins have been described: One being a contractile tail like, similar to tails of T-even phages (R-type) and a non-contractile but flexable tail like similar to those found in λ phages, or those belonging to the family *Siphoviridae* (F-type) (Michel-Briand and Baysse 2002). It has been proposed (Kageyama 1985) that these bacteriocins were once defective phages. Later studies confirmed (Nakayama *et al* 2000) that R and F-type produced by *P. aeruginosa* have common ancestral origins with phages P2And λ respectively. Hence, these defective phages were later adopted and specialized as bacteriocins.

Determining minimal amount of DNA needed for antimicrobial activity for clone 5

Clone 5 was digested with both *Pst*I and *Bgl*II in order to release the insert and also to obtain two fragments. The *Bgl*II - *Pst*I fragment contains ORF2 along with a sizable of ORF1 with a size of 650bp. The *Pst*I – *Bgl*II contains a stretch of sequences with no predicted ORFs but it does contain a small piece of ORF1 with the size of 1150bp. These two subclones was re-ligated back into vector pDA71And transformed into *E. coli* MM294-4. Five colonies were picked from each of the two transformations and screened for inserts

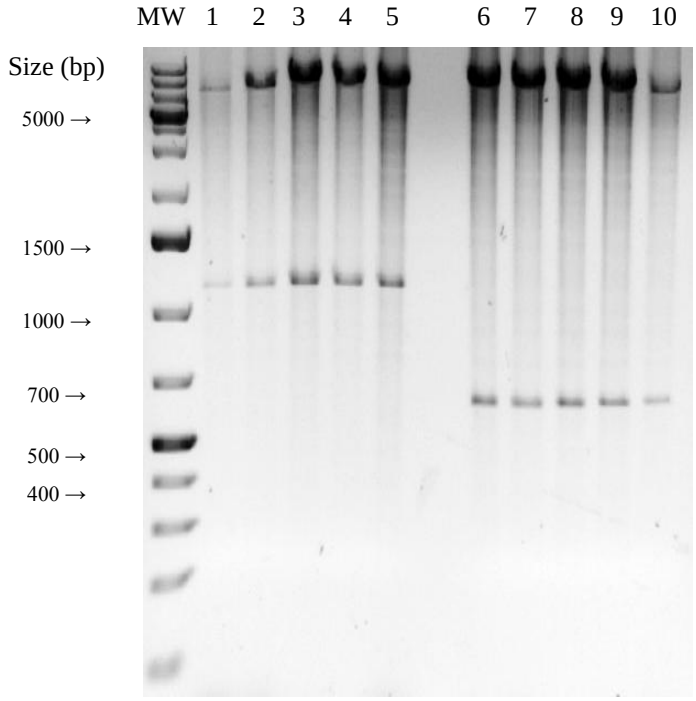



**Figure 3.18** Subclones from clone 5 of Mozambique 1 *Pst*I library: *Pst*I – *Bgl*II are in lanes 1 – 5 and subclone *Bgl*I – *Pst*I are in lanes 6 - 10. Inserts were released after digestion with *Bgl*II and run on a 0.8% agarose gel.

Both fragments was transformed into *R. erythropolis* SQ1 to determine the minimal amount of DNA needed for antimicrobial activity

**Table 3.14** Transformation efficiencies for both subclones of clone 5 from Mozambique 1 *Pst*I library

| Restriction sites at the end of subclone inserts | Insert size (bp) | Transformants/µg DNA |
|---|---|---|
| *Pst*I – *Bgl*II | 1150 | < 6 |
| *Bgl*II – *Pst*I | 850 | 910 |
| pDA71* control | N/A | 800 |

### 3.9.2 Clone 6 of Mozambique 1 *Pst*I Library

Insert of clone 6 were extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below.

ctgcagcagt tgccagatac atcgcggtcg tggtcttacc cacgccgccc ttcgtgtgaa ctaccgagat cgttgtcatg atcggatcgt actcgaatcc tcgaggactag aggtctcgac accccgagga ctcgaaaact cgaggtctcg actactcgat agtcgactaa tcgacaactc gactaatcga tgcaggtcar agacacgcct gtcggggcaa atgcggcggt tgtgactcgt tcgggaggaa gatcagcgcg accgaacaac tccataaacg acgaaacccc ctgttggcgc agggggctgt ctgtcccgaa gggaacgatg acttggacgt ctcgtttcac ggtagcgcac gcctccgaca agcggaagct gcgacacgaa accgtgatgc atttcgtaac tggcttgggc gagctggcgc gtcggccgcg ccgcgatggg tgtctcgcga gcagtggctt aaggacgtga ccgagtggtc gcgatccgac cgattcaccg cctgcggcat ctcgatggac cccatgacct tcatggcgat catgcgggtc atggccgact acgcagacca ctcagacggg cgctgtgtgg ccgtctcccg gtccagaatc gcttccgcag tgggttgtgc gcccggaacg gttacgcggg cctggaggct cctacggytc tccgggtacg ggatcgaggt ccagcgcggc cacggctccg accggacgcc ctcgcagggg tgccggccgt cgatctacca ccttgtgagt aaacgccgtc cggttgtgga caatgttcac ctaccgccga aggcggcagt agttctctct tctcccgtag ggaggtactc accaagggcc gctgcgcggc cgcgaaagat gcctcgaaag ggaaagacaa ccccggccga tcgagctgcag

**Figure 3.19** Sequence from clone 6 of Mozambique 1 *Pst*I library with *Pst*I site highlighted in orange

BLASTn results

The sequence was then analyzed using BLASTn to determine whether it shares any nucleotide sequence similarities with sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| JF704097.1 | Mycobacterium phage Gladiator, complete sequence | 181 | 181 | 17% | 7e-42 |
| JF937092.1 | Mycobacterium phage DaVinci, complete genome | 175 | 175 | 17% | 4e-40 |
| JN049605.1 | Mycobacterium phage EricB, complete genome | 175 | 175 | 17% | 4e-40 |

**Figure 3.20** BLASTn search result from clone 6 of Mozambique 1 *Pst*I library when aligned with known sequences from the database

BLASTx results

The sequence was also analyzed using BLASTx to determine whether it shares any similarities with protein sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| NP_862580.1 | Rep-like protein [Mycobacterium celatum] >gb\|AAD42965.1\|AF14488 | 134 | 134 | 65% | 6e-34 |
| YP_001700714.1 | putative rep protein [Mycobacterium abscessus ATCC 19977] >ref\|YP | 134 | 134 | 43% | 7e-34 |
| CAB43095.1 | putative Rep protein [Mycobacterium fortuitum] >gb\|AAQ24374.1\| Re | 123 | 123 | 43% | 8e-30 |
| YP_001136782.1 | hypothetical protein Mflv_5533 [Mycobacterium gilvum PYR-GCK] >gb | 90.9 | 90.9 | 50% | 5e-18 |

**Figure 3.21** BLASTx search result from clone 6 of Mozambique 1 *Pst*I library when aligned with known protein sequences from the database. These results are shows that although it has a nucleotide match to known mycobacteriophages, the BLASTx results shows that the insert share sequence similarities with several *mycobacterial* proteins.

ORF prediction and construction of restriction map

Open reading frames were predicted using GeneMark and SoftBerry, a given ORF is accepted when consensus is reached between the results from the two programs. Clone 6 has two ORFs, with ORF1 occupying the position between 2 to 79 and ORF2 from 184 to 351. A restriction map of the insert is drawn after analyzing the sequence using NEB cutter showing the location of the restriction sites and those of the two ORFs.

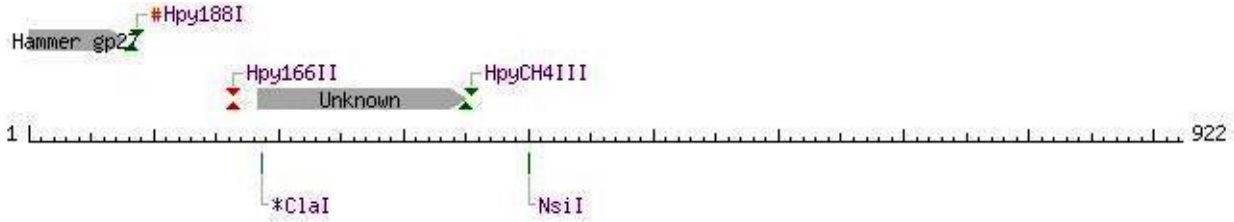



**Figure 3.22** Restriction map of clone 6 of Mozambique 1 *Pst*I library along with the positions of the two predicted ORFs

BLASTp results of Clone 6 ORFs

The two ORFs of clone 6 were analyzed to determine any similarity to proteins in the database. Phage Hammer gp37 is a hypothesised parA protein dealing with chromosome segregation.

BLASTp result of ORF1 from clone 6

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEK08682.1 | gp37 [Mycobacterium phage Hammer] | 82.9 | 82.9 | 100% | 7e-18 |
| AEJ93332.1 | gp37 [Mycobacterium phage EricB] >gb\|AEJ95038.1\| gp37 [Mycobac | 82.9 | 82.9 | 100% | 7e-18 |
| ADB93730.1 | gp37 [Mycobacterium phage RedRock] | 73.6 | 73.6 | 100% | 1e-14 |
| AEK07308.1 | gp35 [Mycobacterium phage PackMan] | 65.5 | 65.5 | 100% | 9e-12 |

**Figure 3.23** BLASTp search result from ORF1 of clone 6 of Mozambique 1 *Pst*I library when aligned with known protein sequences from the database.

BLASTp result of ORF2 from clone 6

BLASTp results from the predicted ORF shows no significant sequence similarity with protein sequences from the database.

During bacterial replication, chromosomal DNA are separated and moved towards the opposite poles of the host. Chromosomal segregation is further aided by homologues of two plasmid partitioning proteins, ParA and ParB, which orientate the origin of replication to towards the poles (Sharpe and Errington 1997; Mohl and Gober 1997). Both proteins are found in plasmids (as SopA) and the bacteriophage P1 which is used for DNA segregation when P1 is replicating itself as a plasmid prophage (Gordon and Wright 2000). ParA is an ATPase that acts as a tracriptional repressor by binding its helix turn helix sequence upstream of a *parAB* operon (Gordon and Wright 2000). Experiments (Easter and Bober 2002) have shown that an overexpression of ParA prevents cell division.

Determining minimal amount of DNA needed for antimicrobial activity for clone 6

Clone 5 was digested with both *Pst*I and *Nsi*I in order to release the insert and also to obtain two fragments. The *Nsi*I - *Pst*I fragment is 522bp with no predicated ORF located in it. The *Pst*I – *Nsi*I contains a stretch of sequences with both ORFs in a short sequence of 400bp.These two subclones was re-ligated back into vector pDA71And transformed into *E. coli* MM294-4. Five colonies were picked from each of the two transformations and screened for inserts

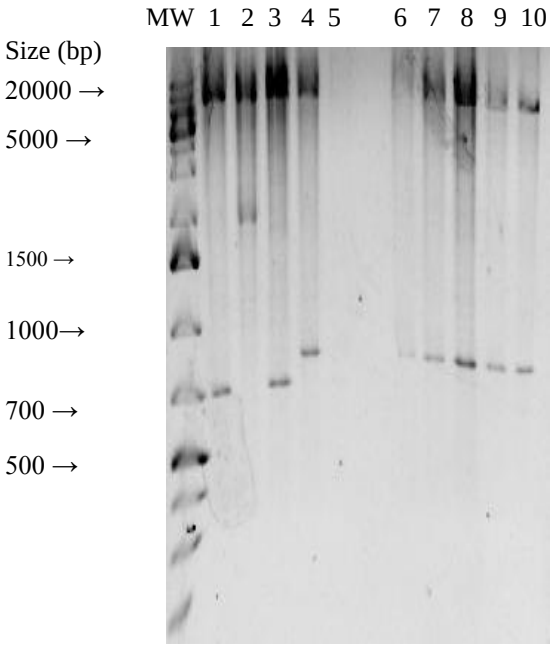


**Figure 3.24** Subclones from clone 6 of Mozambique 1 *Pst*I library: *Pst*I – *Nsi*I are in lanes 1 – 5 and subclone *Nsi*I – *Pst*I are in lanes 6 - 10. Inserts were released after digestion with *Bgl*I and run on a 0.8% agarose gel. Lanes 6 – 10 all contains subclone *Nsi*I – *Pst*I while only lanes 1And 3 contain subclone *Pst*I – *Nsi*I.

Both fragments was transformed into *R. erythropolis* SQ1 to determine if these subclones are able to exhibit antimicrobial activity

**Table 3.15** Transformation efficiencies for both subclones of clone 6 from Mozambique 1 *Pst*I library

| Restriction sites at the end of subclone inserts | Insert size (bp) | Transformants/µg DNA |
|---|---|---|
| *Pst*I – *Nsi*I | 400 | 2 |
| *Nsi*I – *Pst*I | 522 | 5 |
| pDA71* control | N/A | 800 |

### 3.9.3 Clone 22 of Mozambique 3 *Pst*I library

Insert of clone 22 were extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below.

ctgcagccc ccgtastaat mcgttctgca tgatcagctt cctaatcgcc ccgacggcga tccgcacgtc ttgctcgaac acctccccca gctccaccgg cacgacccag gactcgtcgt gagtcccgaa ctggacgagg cggcacgcga acgtgtcgct gtagatgtcc agcccggtgg tctcggtgtc gacggcgagg cagttcarat gagcccggat gaagtcgcgg aagccgtcca gatcctctgg gtgctcaac gacgttgatg gtgacgaggt ctcctccgac ctcatgccgt agctcgatca tcttctctcc tagctgtatg ggtttcgctc gacgctggcc ttcgagtagt ggcggctctg gtggtcggtg tggacgaccg ggccgcaaac tcccttgtgg gccttgccat ctgcgtagaa caggacgccg atcaggttgt gctgccacag agacagcagc ccgaactcac ccatgtccaa ggtggctgag gtcggtggtc cgcgtccgt atcggcaacg tgttctcgtc cacatcagcg atctcgctca tgaggttgat cgccttgcgg cgcaggttga ctccctggac atcgagcggc agagcggggt cgatcggctc cgcgatcatg acccggattc ggtctgccat cagtggtaga tccccccggac ggtgcgcgag atcgtggccg ggttcacgcc gtagttgcgg gcgagatcct ttctgcttcg cgccgccgta gtaggcgtca cgaatgtcct tggcctcggt ctcggtgagc ttgcggcggt tcggccggct cgggccggtg ggagcctgcg gctcgccctt gacgaacgcc tctccgaagc atcgcttggc ggtctcgagc tgccgcttca ggtcgtggtt ggaccggatg gcgctgttcg cgacagcccg acagcttcac attgggccgc atgaaagctc tgcgttcttc gcccgcacgg tgatcagctc cccgaacggt ctggttgttc tcgggcctcc agcgaggcgc tgaactcctt ctccgctgcc agctcagctc gcartgcttt tcttcgtctt ggtatgccat cagttgggcc tcctgcctga atttggctcg tactcttc-t tcggtcaatg tagtagaagt cgatgacctt gtcccagttg aaagttcggg acgtgccgtc atcgaacgcg atgatcagga caccctcggt ggtgtcgagg atcggctcac ccgcgatgat gtgataccgg tcctcgaggt tcgccacggt cgctcgtcgt gcttgcatgt cagcctccgt acggctcgtc ggggatgtcc tggtaggagt tgggagcgat ctcccggagc tgccgcagca gttcccctgc cagttcacgg atttcggcat ccgcggcctc gtgccagcgg gccttgatga cgttgcgcca cgcccggtgg ttgccggtg acgaccatcg gtgagttcgt catgttcggc aggaccgcac gggctgcctc acgaacctgc ttgcggggca agcccttcgc cgtgaggtgg tccacgatcg tcccgtaggt cgtcgtggct agatcccaga tgagctgcag

**Figure 3.25** Sequence from clone 22 of Mozambique 3 *Pst*I library with *Pst*I site highlighted in orange.

BLASTn results

The sequence was then analyzed using BLASTn to determine whether it shares any nucleotide sequence similarities with sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| JN408461.1 | Mycobacterium phage Trixie, complete genome | 514 | 1030 | 57% | 5e-142 | 88% |
| Z18946.1 | Mycobacteriophage L5 complete genome | 418 | 922 | 52% | 5e-113 | 88% |
| AF022214.1 | Mycobacteriophage D29, complete genome | 405 | 719 | 73% | 6e-109 | 82% |
| JN408460.1 | Mycobacterium phage Turbido, complete genome | 385 | 1066 | 60% | 4e-103 | 90% |

**Figure 3.26** BLASTn search result from clone 22 of Mozambique 3 *Pst*I library when aligned with known sequences from the database

BLASTx results

The sequence was also analyzed using BLASTx to determine whether it shares any similarities with protein sequences in the database. No sequence similarity was found

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| AEJ93473.1 | gp44 [Mycobacterium phage Benedict] | 158 | 158 | 17% | 3e-42 | 82% |
| AEL17817.1 | gp50 [Mycobacterium phage Turbido] | 155 | 155 | 18% | 2e-41 | 77% |
| AEL17912.1 | gp47 [Mycobacterium phage Trixie] | 155 | 155 | 18% | 3e-41 | 80% |
| AEJ94558.1 | gp46 [Mycobacterium phage Backyardigan] >gb\|AEL19913.1\| gp47 [M | 152 | 152 | 17% | 3e-40 | 79% |

**Figure 3.27** BLASTx search result from clone 22 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

ORF prediction and construction of restriction map

Open reading frames were predicted using both GeneMark and SoftBerry, an ORF is accepted when there is a consensus between the results from the two programs. Clone 22 has four ORFs, with ORF1 occupying the position from 1 to 310; ORF2 from 319 to 483, overlapping with ORF3, which starts from 476 to 658. ORF4 lies further away, occupying a position from 1298 to 1596. A restriction map of the insert, along with the location of the predicted ORF and the restriction sites is drawn after analyzing the sequence using NEB cutter.

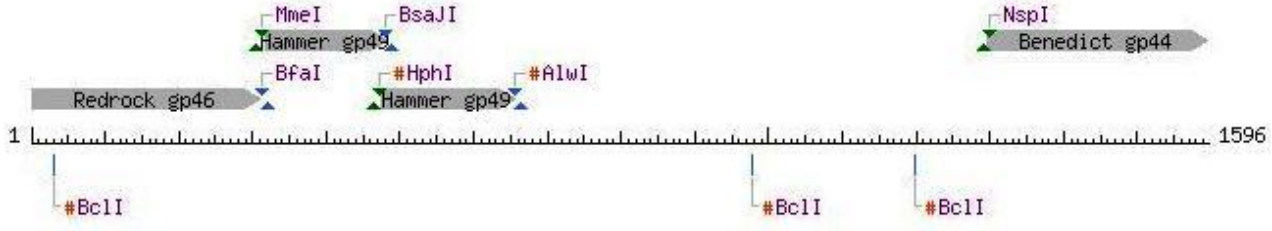



**Figure 3.28** Restriction map of clone 22 of Mozambique 3 *Pst*I library along with the positions of the four predicted ORFs

BLASTp results of Clone 22 ORFs

The four ORFs of clone 22 were analyzed to determine any similarity to proteins in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| ADB93739.1 | gp46 [Mycobacterium phage RedRock] | 176 | 176 | 93% | 3e-51 |
| AEL17915.1 | gp44 [Mycobacterium phage Trixie] | 175 | 175 | 93% | 9e-51 |
| AEL17822.1 | gp45 [Mycobacterium phage Turbido] | 174 | 174 | 93% | 3e-50 |
| NP_039708.1 | predicted 66.2Kd protein [Mycobacterium phage L5] >sp\|Q05254.1\|D | 170 | 170 | 93% | 5e-49 |
| YP_001994860.1 | gp43 [Mycobacterium phage Pukovnik] >gb\|ACE79969.1\| gp43 [Mycc | 154 | 154 | 93% | 8e-43 |
| NP_046860.1 | DNA polymerase [Mycobacterium phage D29] >sp\|O64235.1\|DPOL_BP | 149 | 149 | 93% | 4e-41 |

ORF1

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEK08694.1 | gp49 [Mycobacterium phage Hammer] | 102 | 102 | 98% | 5e-28 |
| AEJ95088.1 | gp46 [Mycobacterium phage Gladiator] | 91.3 | 91.3 | 98% | 1e-23 |
| AEJ93385.1 | gp47 [Mycobacterium phage EricB] | 91.3 | 91.3 | 98% | 2e-23 |
| AEK08490.1 | gp47 [Mycobacterium phage DaVinci] | 89.0 | 89.0 | 98% | 1e-22 |

ORF2

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEK08694.1 | gp49 [Mycobacterium phage Hammer] | 69.7 | 69.7 | 83% | 4e-15 |
| AEJ95088.1 | gp46 [Mycobacterium phage Gladiator] | 68.9 | 68.9 | 83% | 8e-15 |
| AEK08490.1 | gp47 [Mycobacterium phage DaVinci] | 68.2 | 68.2 | 83% | 1e-14 |
| AEJ93385.1 | gp47 [Mycobacterium phage EricB] | 65.9 | 65.9 | 83% | 1e-13 |
| AEL17821.1 | gp46 [Mycobacterium phage Turbido] | 58.9 | 58.9 | 98% | 6e-11 |

ORF3

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEJ93473.1 | gp44 [Mycobacterium phage Benedict] | 160 | 160 | 92% | 1e-48 |
| AEL17817.1 | gp50 [Mycobacterium phage Turbido] | 159 | 159 | 100% | 5e-48 |
| AEL17912.1 | gp47 [Mycobacterium phage Trixie] | 158 | 158 | 98% | 2e-47 |
| NP_039712.1 | predicted 27.8Kd protein [Mycobacterium phage L5] >sp\|Q05259.1\|T | 155 | 155 | 88% | 2e-46 |
| AEK32660.1 | gp41 [Mycobacterium phage George] | 154 | 154 | 93% | 3e-46 |
| AEJ94558.1 | gp46 [Mycobacterium phage Backyardigan] >gb\|AEL19913.1\| gp47 [M | 154 | 154 | 94% | 4e-46 |
| ADB93744.1 | gp51 [Mycobacterium phage RedRock] | 153 | 153 | 93% | 9e-46 |

ORF4

**Figure 3.29** BLASTp search result of ORF1, 2, 3 and 4 from clone 22 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

Redrock gp46 is a DNA polymerase I with both a 3' – 5' exonuclease and a polymerase domain. It is involved in the replication in both bacteria (*E. coli*) and phages (T4) (Moser *et al* 1997) and is vital in steps such as proofreading, repair, DNA processing and stability. In *E. coli*, the exonuclease domain is comprised of three residues (Asp-355, Asp-424 and Asp-501) (Reha-Krantz and Nonay 1993). Although the 3D structure of both bacterial and phage of this domain are structurally the same, their protein sequence shares very little similarity, and mutations at this domain results in limited proofreading activity (Moser *et al* 1997).

In phage Φ29, other than 3' – 5' activity and DNA polymerization, polymerase of this phage can also perform polymerization reversals and it can also replicates phage DNA without the use of Okazaki fragments or DNA helicases (Blanco and Salas 1996), making it highly processive (>70 kb) during phage replication.

It could be suggested that such high processivity during replication ensures the phage will have numerous copies of its genome shortly after infection. This could link to high processivity of transcription of certain phages such as T7. In phage T7, transcription of phage genes are 5 fold faster than those of host genes, resulting in long lines of phage mRNA (Marchand *et al* 2001). Other phage proteins, such as RNase III and gp0.7, will also work in conjunction with phage mRNA to enhance its stability and assist in translation (Roucourt and Lavigne 2009)

Increase in the replication of phage genome could also influence the replication of the host. Studies have shown that some phages redirect many of the host replication machinery in order to replicate its genetic material (Friedman *et al* 1984). However, other studies (Powell *et al* 1992) have also shown that c6A, a phage that infects *Lactococcus lactis*, rapidly degrades host cell DNA and incorporates them into phage progenies. The high processivity exhibited by Φ29 could help ensure rapid degradation of host DNA as seen in phage c6A.

Benedict gp44 is a thymidylate synthase complementary protein (ThyX/Thy1), which has the same function as thymidylate synthase (TS) although it shares no homology with it. TS is encoded by the gene ThyA, which is responsible for the *de novo* synthesis of 2'-deoxythymidine 5'-monophosphate through the methylation of 2'-deoxyuridine 5'-monophosphate, thereby making the enzyme essential for DNA replication (Sampathkumar *et al* 2005). Although it has been thought that TS was responsible for this in all organisms, genomic studies into several bacterial and archeal genomes has not found genes coding for TS (Myllykallio 2003). This leads to the discovery of another enzyme (ThyX/Thy1) that carries out the same function (Myllykallio 2002). Due to its importance in DNA replication, it is a conserved sequence in pathogens such as *Helicobacter pylori*, and in mycobacterium such as *M. leprae* and *M. tuberculosis*, and has widely considered to be a target for any future antimicrobial drugs.

*M. leprae* and *M. tuberculosis* has both ThyA and ThyX gene, and has been hypothesized that the cell will use either one depending on different growth conditions (Purwantini *et al* 1997). ThyX is also reported to be in mycobacteriophages D29 and D5 (Myllykallio 2002) and it could be suggested that these genes found it's way into the phage genome either through packaging or through co-evolution with their bacterial hosts, especially a viruses that seeks to increase its replication speed and efficiency. This is true for viruses that has no access to thymidylate in the early stages of the infection, with an example of this being Paramecium bursaria chlorella virus-1 (Graziani *et al* 2004).

### 3.9.4 Clone 29 of Mozambique 3 *Pst*I library

Insert of clone 29 was extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below.

ctgcaggcsg gtkgtgggag gagattccgc gagccatctt gaggatcctct cctgatggcc tacgtcaacg gggccatcat cctctggggc ttcttcacct ggctcagcct ggtgctcgat caccgagagg agcgtgatcg tgcccgcgaa gcggaagccg ccggccaggc cgaaagctcc taagcagtgc gtcgattgcg ttgctgaggg gatcacgtcg aagcgcaaga ccccacaccc cggtcctcgg tgcgtaacgc actggcgggc caagaagcgc gagcgcagct cgggcgcgtg gggcgcgaga atccaagcca cgtatggcat tacggctgac gagtactggg cgatctacga gttccagggc ggtcgctgct acatctgcca gcgtgcgaac gggaaggtga aacgactat cggtagacca tgaccacaag acgggcatca tccgagggct gctctgcac gatgtgcaac aagtacaccc tgggctgggc gagggactgc atcgagttct tcaagcgggc catcgcgtac ctgacgaatc ctcctgcggt gcaagtcatc ggggagcgca tcgctcccat cgaggctgaa aggctgaccc gaacttgaca gccaccagcc cgatcgcagt agccatccag cggtactacc cggactggga agcaccgccc gatcacaacg agtggaacaa gtgcctgtgc ccgttccacg gcgacgaaac gccgtctgcc gcagtcagtt acgacctgcag

**Figure 3.30** Sequence from clone 29 of Mozambique 3 *Pst*I library with *Pst*I site highlighted in orange.

BLASTn results

The sequence was then analyzed using BLASTn to determine whether it shares any nucleotide sequence similarities with sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| AF022214.1 | Mycobacteriophage D29, complete genome | 439 | 482 | 57% | 1e-119 | 96% |
| GU339467.1 | Mycobacteriophage RedRock, complete genome | 426 | 600 | 69% | 1e-115 | 92% |
| JN408461.1 | Mycobacterium phage Trixie, complete genome | 403 | 548 | 68% | 1e-108 | 89% |
| DQ398043.1 | Mycobacteriophage Che12, complete genome | 378 | 542 | 65% | 4e-101 | 91% |
| JN408460.1 | Mycobacterium phage Turbido, complete genome | 355 | 534 | 57% | 3e-94 | 89% |
| JN049605.1 | Mycobacterium phage EricB, complete genome | 347 | 509 | 69% | 7e-92 | 90% |

**Figure 3.31** BLASTn search result from clone 29 of Mozambique 3 *Pst*I library when aligned with known sequences from the database

BLASTx results

The sequence was also analyzed using BLASTx to determine whether it shares any similarities with protein sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| ADB93756.1 | gp63 [Mycobacterium phage RedRock] | 124 | 124 | 54% | 7e-33 | 91% |
| AEL17900.1 | gp60 [Mycobacterium phage Trixie] | 123 | 123 | 54% | 1e-32 | 86% |

**Figure 3.32** BLASTx search result from clone 29 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database

ORF prediction and construction of restriction map

Open reading frames were predicted using GeneMark and SoftBerry, an ORF is accepted when a consensus result is reached from the two programs. Clone 29 has three ORFs, with ORF1 occupying the position from 56 to 184; ORF2 from 471 to 617, and ORF3, which starts from 614 to 760. A restriction map of the insert is also drawn after analyzing the sequence using NEB cutter. The positions of the three OFRs are also shown on the map.

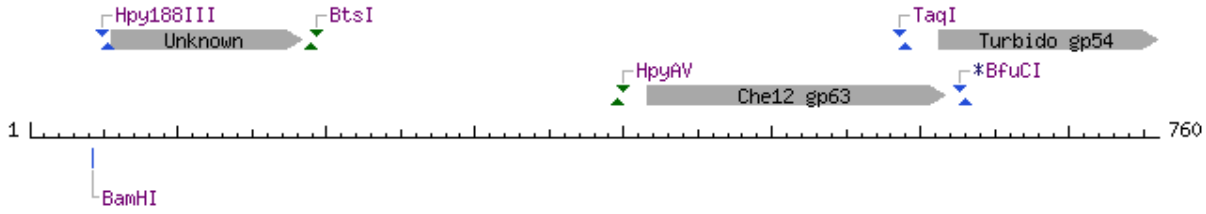



**Figure 3.33** Restriction map of clone 29 of Mozambique 3 *Pst*I library along with the positions of the three predicted ORFs

BLASTp results of Clone 29 ORFs

The three ORFs of clone 29 were analyzed to determine any similarity to proteins in the database. ORF1 did not show any significant similarity.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| YP_655642.1 | gp63 [Mycobacterium phage Che12] >gb\|ABE67382.1\| gp63 [Mycoba | 97.4 | 97.4 | 100% | 1e-25 |
| NP_039723.1 | predicted 18.9Kd protein [Mycobacterium phage L5] >sp\|Q05272.1\|V | 94.7 | 94.7 | 100% | 2e-24 |
| ADB93756.1 | gp63 [Mycobacterium phage RedRock] | 92.8 | 92.8 | 93% | 7e-24 |

ORF2

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEL17806.1 | gp61 [Mycobacterium phage Turbido] | 96.7 | 96.7 | 97% | 2e-25 |
| AEJ95100.1 | gp59 [Mycobacterium phage Gladiator] >gb\|AEK08751.1\| gp62 [Mycc | 94.0 | 94.0 | 91% | 2e-24 |
| AEL17924.1 | gp58 [Mycobacterium phage Trixie] | 93.2 | 93.2 | 100% | 3e-24 |
| YP_001994875.1 | gp58 [Mycobacterium phage Pukovnik] >gb\|ACE79984.1\| gp58 [Mycc | 92.4 | 92.4 | 95% | 7e-24 |

ORF3

**Figure 3.34** BLASTp search result of ORF 2 and3 from clone 29 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

Recognition of DNA junctions and resolving them is essential to DNA recombination and repair, and enzymes that carry out this function have been seen in all forms of life (Raaijmakers 1999). In mycobacteriophage Che12, this function is carried out by endonucelase VII encoded by gp63, to which ORF 2 of clone 29 shares similarity. This enzyme has been extensively studied in the phage T4, which is expressed at early and late stages of infection and is involved in solving junctions (such as Holiday junctions) (Mizuuchi 1982), mismatch repair, and resolving branch points prior to packaging into phage heads (branching is also a feature in DNA recombination) (Solaro *et al* 1993; Grebenshchikova *et al* 1994). As such, T4 phages that carry a mutation of this gene will result in defective packaging with heads containing highly branched DNA and also reduced DNA replication (Raaijmakers 1999, Miller *et al* 2003).

Like other resolvases, it acts as a dimer, and nicks branched DNA on both strands while working with DNA polymerase and ligase to repair DNA. Although it carries out the basic responsibilities of a resolvase, endonuclease VII is also involved in resolving Y-junctions, heteroduplex loops, single strand overhangs and single base mismatches, all this suggests that endonucelase VII is descended from an ancient class of repair enzymes (Kemper 1997, Raaijmakers 1999).

It is important to note that a resolvase found in T4, is also found in other bacteriophages such as T7, RusA, and Rap (Sharples 2001). This is due to recombination being an important process of both phage replication and evolution. During infection, recombination is needed to initiate DNA replication in T4 (Mosig 1998). It is also important to repair any damage DNA in order to increase the efficiency of replication to maximize the number to phage progeny. Recombination also plays a very important role in phage evolution, as seen by the presence of many recombination enzymes encoded by phages (Sharples 2001). This is essential for any genetic exchanges between phages and acquiring new beneficial genes, further enhancing the mosaic makeup already seen.

For ORF3, it bears similarity to gp61 of phage Turbido, a DNA primase. This enzyme produces primers during DNA replication by synthesizing oligoribonucleotides *de novo* on single strand DNA. In phages T7 and T4, DNA primase recognizes a trinucleotide sequence of 3'-CTG-5' and 3'-T(C/T)G-5' respectively and begin synthesis of primers at these sites for DNA polymerase to produce a nascent leading and lagging strand (Kusakabe and Richardson 1996). All primase have a potential metal binding site, and for both phages, this site is on the N-terminus of the primase and has been shown to be a zinc motif (Mendelman *et al* 1994). This motif is important in two ways: To maintain a stable bind between the DNA and the primase in such a way that the catalytic domain can synthesize the primer and to recognize the primer recognition site and synthesizes a primer of 4 – 5 bp long (Kuchta and Stengel 2010). In addition, to stimulate activity, primase are found to be coupled their activity with another enzyme such as helicase. In T7, a single primase protein can carry out both primase and helicase activity (Frick and Richardson 2001)

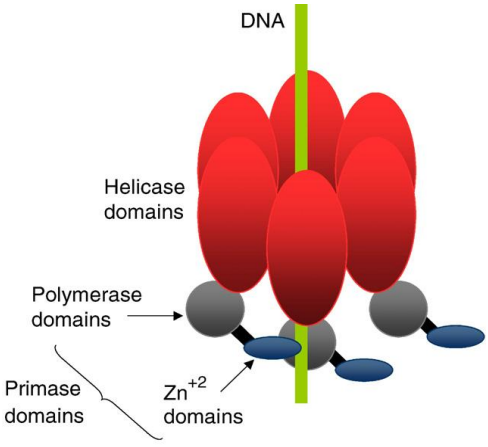



**Figure 3.35** Model of T7 primase-helicase hexamer showing zinc binding domain (Adopted from Kuchta and Stengel 2010)

Due to its essential role in DNA replication, primase, like TS, are now considered to be an important antimicrobial target. Some have been synthesized using oligonucleotide analogues (Kutcha *et al* 1992) while others exists naturally (Wilson and Trauner 2007).

### 3.9.5 Clone 42 of Mozambique 3 *Pst*I library

Insert of clone 42 were extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below.

ctgcaggat cgtggccaga gtcgctgtcg gacgatcagt tcaacggtcg caacctctac tcgtcggaca acgtcaagga cgctctcgac aacgtgtcga ccaatcggta cctggtggac atcttgccga gggctctgga ggcgctggcc gagaagaacg agggccaggc cgaggcgatc aggagccgct acgaggacgg gatcgtgccg tccaagaagg acggcgcagc gatgctgctg tctcgggcgg tcaagtcgct gaccgaaggc acgttcaaca ttcatcgcga tcacctccgg tgtcgacgct gacggcaaac gtcaccgagg gtccggggag ccggcactca gtgttcccgg agacccgccc caccagcggc gggcactccg atccgacggc cgacatcgcg atcctgctga tcgagcatcc cgagctgcgc gacgagtacc tcgcggagcc gtcgctgccc gagttcttgg gagggaggtg ctatgcacag cctgcttgat cccacgttca acgggatgcc gggttcggag atgtaccggg gcgaagtgtt cccagagttg ttccccggca agcgaatgat gctcgagaac tggtcgcaag acgacctaga gcagtacgtc ggaggggtct tcacccccgg atatggaatg agaggagctg cgtgacagac accgaaatac cttgggggcc aacgggtgaa ctcgtctaca atcgtactta cgcacgaatc aagcctgatg gaacgcgtga gacgtggcca gagacagttc ggcgagtcgt ggacggcaat cttgcgcttg tgcctgagcg ataccaccag gacggggagc gcgaacagct cgtccgactg atggagcagt tcaagatcct gcccgccggc cggcacctgt gggccagcgg cgtgaagaac gcccagcacc tgttcaactg ctgggtgtcg ggatggcccg aggacatctc ggaccacttc cagttcactt tcatgcgcct gatggagggc gggggaagtc ggaagcgaac tactccaacc actacctcga gcactactcg cacgtcgtgc atcccctccg ggtcgagatc gtctgcgacc cagagcatgg tcgacctacc aggcgatgaa ggacgctggc atcctgtcga agcggtacga ctccgaatgg gcaggcgcgt tcccgatcga ggactcccgc gagggctggg ccaacgccct ggtcgacctg atcgacacgc actaccgacg cgacacggtcc acttcagccg ggtctacgac gtgagtcgta tccggcctgc cggggccaag ctcaagackt tcggtggatt cgccagcggg ccgctgccgt tcgctcagat gctgcag

**Figure 3.36** Sequence from clone 42 of Mozambique 3 *Pst*I library with *Pst*I site highlighted in orange.

BLASTn results

The sequence was then analyzed using BLASTn to determine whether it shares any nucleotide sequence similarities with sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| DQ398043.1 | Mycobacteriophage Che12, complete genome | 556 | 881 | 83% | 1e-154 | 81% |
| Z18946.1 | Mycobacteriophage L5 complete genome | 498 | 827 | 81% | 3e-137 | 80% |
| JF704110.1 | Mycobacterium phage PackMan, complete sequence | 393 | 455 | 53% | 2e-105 | 86% |
| EU744250.1 | Mycobacterium phage Pukovnik, complete genome | 368 | 663 | 79% | 8e-98 | 87% |
| JN408461.1 | Mycobacterium phage Trixie, complete genome | 346 | 649 | 94% | 2e-91 | 78% |

**Figure 3.37** BLASTn search result from clone 42 of Mozambique 3 *Pst*I library when aligned with known sequences from the database.

BLASTx results

The sequence was also analyzed using BLASTx to determine whether it shares similarities with protein sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| YP_655632.1 | gp53 [Mycobacterium phage Che12] >gb\|ABE67372.1\| gp53 [Mycoba | 209 | 360 | 49% | 4e-84 | 88% |
| NP_039714.1 | predicted 76.3Kd protein [Mycobacterium phage L5] >sp\|Q05262.1\|V | 210 | 351 | 49% | 2e-81 | 85% |

**Figure 3.38** BLASTn search result from clone 42 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

ORF prediction and construction of restriction map

Open reading frames were predicted using both GeneMark and SoftBerry, an ORF is accepted when there is a consensus between the results from the two programs. Clone 42 has two ORFs, with ORF1 occupying the position from 492 to 674; ORF2 from 1125 to 1367. A restriction map of the insert is drawn after analyzing the sequence using NEB cutter. The location of the three ORFs is also plotted on the map. This insert does not have any restriction sites that would allowing for subcloning in pDA71, with *Sau3*AI and *Taq*I being the only two endonuclease able to digest this insert.

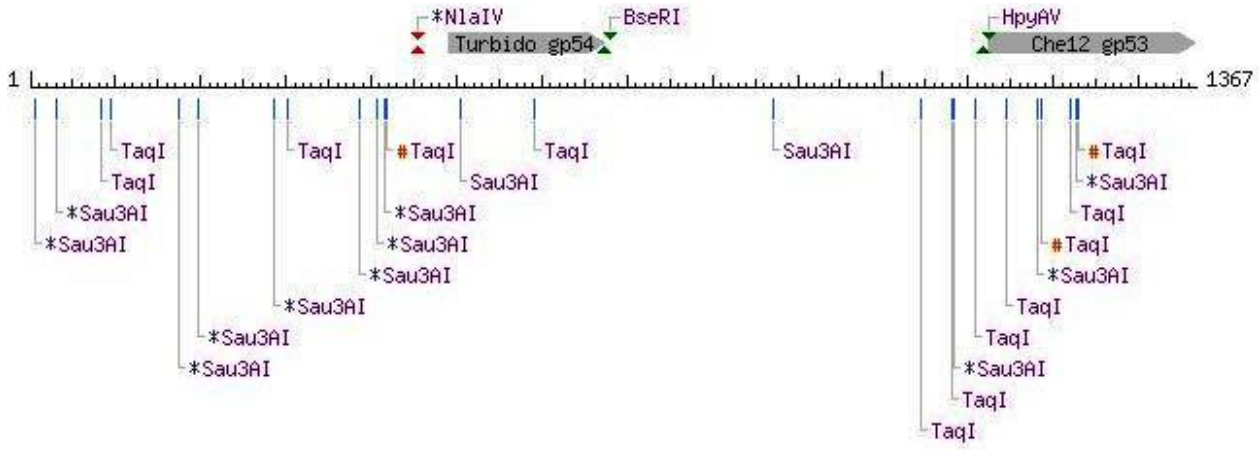



**Figure 3.39** Restriction map of clone 42 of Mozambique 3 *Pst*I library along with the positions of the two predicted ORFs

BLASTp results of Clone 42 ORFs

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEL17813.1 | gp54 [Mycobacterium phage Turbido] | 107 | 107 | 100% | 2e-30 |
| NP_046867.1 | gp52 [Mycobacterium phage D29] >sp\|O64242.1\|VG52_BPMD2 RecNa | 102 | 102 | 95% | 1e-28 |
| AEL17908.1 | gp51 [Mycobacterium phage Trixie] | 98.2 | 98.2 | 91% | 8e-27 |
| YP_655633.1 | gp54 [Mycobacterium phage Che12] >gb\|ABE67373.1\| gp54 [Mycoba | 97.1 | 97.1 | 96% | 3e-26 |
| NP_039716.1 | predicted 7.0Kd protein [Mycobacterium phage L5] >sp\|Q05268.1\|VG | 96.3 | 96.3 | 100% | 5e-26 |

ORF1

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| YP_655632.1 | gp53 [Mycobacterium phage Che12] >gb\|ABE67372.1\| gp53 [Mycoba | 150 | 150 | 100% | 1e-41 |
| AEK07323.1 | gp50 [Mycobacterium phage PackMan] | 137 | 137 | 100% | 4e-37 |
| NP_039714.1 | predicted 76.3Kd protein [Mycobacterium phage L5] >sp\|Q05262.1\|V | 135 | 135 | 100% | 3e-36 |
| YP_003358753.1 | gp50 [Mycobacterium phage Peaches] >gb\|ACU41801.1\| gp50 [Mycc | 133 | 133 | 98% | 1e-35 |
| ADL71310.1 | gp50 [Mycobacterium phage Eagle] | 133 | 133 | 98% | 1e-35 |
| AEM05935.1 | gp50 [Mycobacterium phage TiroTheta9] | 133 | 133 | 98% | 1e-35 |
| AEF57379.1 | gp50 [Mycobacterium phage Shaka] >gb\|AEL19825.1\| gp50 [Mycoba | 133 | 133 | 98% | 1e-35 |

ORF2

**Figure 3.40** BLASTp search result of ORF 1 and 2 from clone 42 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

Ribonucleotide reductase (RNR) is essential to DNA synthesis and repair since it converts the 4 standard ribonucleotides to their 2'-deoxyribonleotide counterparts (Kolberg *et al* 2003). It is important to note that one such product, 5'-di/triphospho-2'-deoxyuridine, is to be further processed and made into thymidine by TS, another enzyme essential to DNA synthesis. The importance of RNR is evident in its presence in every living organism (Jordan and Riechard 1998) and it could be suggested that certain viruses, including phages carry their own copy of the gene (and other genes that are indispensible to DNA synthesis) in order to ensure faster proliferation of viral genetic material in host cells (Kolberg *et al* 2003).

RNR is divided into three classes based on the different metal co-factors that they utilize (Reichard 1993). However, all three classes share a cysteine residue which is converted into a thiyl radical that subtracts a hydrogen atom from its substrates (Eklund *et al* 2001). This radical is located in a protein loop that is shared by all three classes. This shows that all three classes of RNR share common ancestry with an ancient reductase (Kolberg *et al* 2003).

In phage T4, an anaerobic RNR, *nrdD*, has amino sequences similar to *E. coli* pyruvate formate-lyase (Pfl) (Young *et al* 1994). Both enzymes require radicals for their catalytic activity (in the case of Pfl, it synthesizes formate and acetyl-CoA from cleaving pyruvate) and both require an activating protein to function (for T4, it is gp55.9, for Pfl, it is Pfl activase) thus suggesting that both share common ancestry with an ancient anaerobic enzyme (Young *et al* 1994).

### 3.9.6 Clone 57 of Mozambique 3 *Pst*I library

Insert of clone 57 were extracted from a 0.6% agarose gel, purified, cloned into the *Pst*I site of vector pUC18/19 vector and sequenced. The sequence is shown below

ctgacaggcc ctggcgcssg twmattgraa tcccagcggg accgggtcga cgtatcgct gagagaacma caccacgctg aaggcttctc mggtggcggt gacgctcagg ctcgacccag caccgacctg ctggttctcg atgtagaacg tcgctgaggc gtgctccagc acgctctcgt ggccgacctc gaggatgtgg ttgaggtagt cctcgttctc cgctgtagcc gggttgggcc ggtcgaatga ccggtagcag ttccggcctg cgaactcggc tagctcgtcg gcatcgaagt cgccgaacgg gccgtcatcg tcgggctcgt tgaagatgtc gggctcgaac ccgatctccc gcagcgcatc cgggtcaacc tcggtggctg caatcagttt gactttcata ctctccgctc aragkgrgga ggaggggccg aagcccctcc cccggtggat gtcaagtacc gcgactactt cttgtcgttg aggtactgcg ccttgcactg ctggtcgcga ggtgcggtgc aggaraacag cttgtaggcg ttgcccgcct tgctgacgcc gctcttgaac accatctcgc cgtgctcaca gaaccgcttc tccccgttcg gagcctgctg cgccgcctgc ggcgcacgcg actgctgctg accgccgccg cccgcgttgc cggccggctt cgatyccgct gcgcggaggt gttcaccttc gccagaacgt cggccgggtc cagagccctt cacgaccacc cacgggtcgc tgtaagcgcc ccgcgaactt gaacgtggcc gacacccccg tcagtggagt gctggaccga caccgaatcg accgcagccg ccgaggcggt ggtcgctacc ggcgcggccg gctcakcctg agcgacgggc tcgggctcct gcgagggcgt accccaaggg ycttcgtagg acaaaattgc ttacctttc acttaawggg gacatgcgcc gattggcgca ctcttcawtc gacammccgt tcttcrgacg ggctttttg gcagcagcag attcgtactg ctgcttggtg attcgctcgt agggagcctg cgggaagctc gcctccggga agatcgtgga gcccttgatc agccccgcga acttctgcag

**Figure 3.41** Sequence of clone 57 from Mozambique 3 *Pst*I library with *Pst*I site highlighted in orange.

BLASTn results

The sequence was then analyzed using BLASTn to determine whether it shares any nucleotide sequence similarities with sequences in the database. In the case of clone 57, segments of the insert is found to be similar to nucleotide sequences of mycobacteriophage Che12And Packman

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| YP_655632.1 | gp53 [Mycobacterium phage Che12] >gb\|ABE67372.1\| gp53 [Mycoba | 150 | 150 | 100% | 1e-41 |
| AEK07323.1 | gp50 [Mycobacterium phage PackMan] | 137 | 137 | 100% | 4e-37 |
| NP_039714.1 | predicted 76.3Kd protein [Mycobacterium phage L5] >sp\|Q05262.1\|V | 135 | 135 | 100% | 3e-36 |
| YP_003358753.1 | gp50 [Mycobacterium phage Peaches] >gb\|ACU41801.1\| gp50 [Mycc | 133 | 133 | 98% | 1e-35 |
| ADL71310.1 | gp50 [Mycobacterium phage Eagle] | 133 | 133 | 98% | 1e-35 |
| AEM05935.1 | gp50 [Mycobacterium phage TiroTheta9] | 133 | 133 | 98% | 1e-35 |
| AEF57379.1 | gp50 [Mycobacterium phage Shaka] >gb\|AEL19825.1\| gp50 [Mycoba | 133 | 133 | 98% | 1e-35 |

**Figure 3.42** BLASTn search result from clone 57 of Mozambique 3 *Pst*I library when aligned with known sequences from the database.

BLASTx results

The sequence was also analyzed using BLASTx to determine whether it shares any similarities with protein sequences in the database.

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value | Max ident |
|---|---|---|---|---|---|---|
| AEJ93380.1 | gp52 [Mycobacterium phage EricB] >gb\|AEK08494.1\| gp51 [Mycobac | 94.4 | 94.4 | 12% | 4e-20 | 89% |
| AEJ95084.1 | gp50 [Mycobacterium phage Gladiator] | 94.4 | 94.4 | 12% | 1e-19 | 89% |
| AEK08698.1 | gp53 [Mycobacterium phage Hammer] | 92.8 | 92.8 | 12% | 5e-19 | 86% |
| YP_655631.1 | gp52 [Mycobacterium phage Che12] >gb\|ABE67371.1\| gp52 [Mycoba | 90.5 | 90.5 | 12% | 1e-18 | 84% |

**Figure 3.43** BLASTx search result from clone 57 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database.

ORF prediction and construction of restriction map

Open reading frames were predicted using both GeneMark and SoftBerry, an ORF is accepted when there is a consensus between the results from the two programs. Clone 57 has two ORFs, with ORF1 occupying the position from 3 to 408; ORF2 from 474 to 758. A restriction map of the insert is also drawn after analyzing the sequence using NEB cutter. The positions of the three ORFs are also plotted on the map.

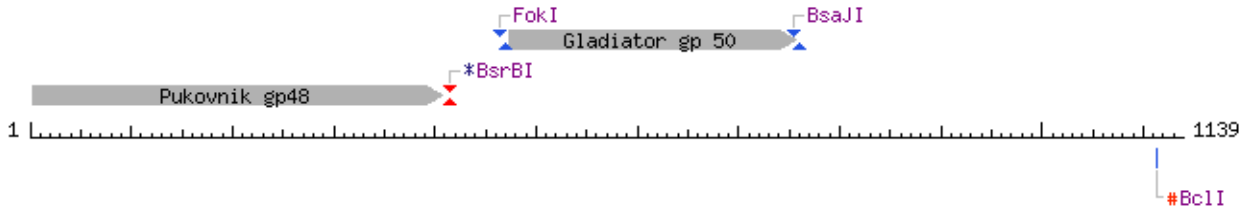



**Figure 3.44** Restriction map of clone 57 of Mozambique 3 *Pst*I library along with the positions of the two predicted ORFs

BLASTp results of Clone 57 ORFs

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| YP_001994865.1 | gp48 [Mycobacterium phage Pukovnik] >gb\|ACE79974.1\| gp48 [Mycc | 186 | 186 | 95% | 5e-58 |
| AEL17817.1 | gp50 [Mycobacterium phage Turbido] | 179 | 179 | 95% | 3e-55 |
| AEJ92066.1 | gp49 [Mycobacterium phage HelDan] | 172 | 172 | 94% | 9e-53 |
| AEK07414.1 | gp49 [Mycobacterium phage Rockstar] | 172 | 172 | 94% | 1e-52 |
| AEL17912.1 | gp47 [Mycobacterium phage Trixie] | 172 | 172 | 95% | 1e-52 |
| ADB93744.1 | gp51 [Mycobacterium phage RedRock] | 171 | 171 | 95% | 3e-52 |
| NP_039712.1 | predicted 27.8Kd protein [Mycobacterium phage L5] >sp\|Q05259.1\|T | 167 | 167 | 94% | 2e-50 |
| AEJ95085.1 | gp49 [Mycobacterium phage Gladiator] | 166 | 166 | 95% | 4e-50 |

ORF1

**Sequences producing significant alignments:**

| Accession | Description | Max score | Total score | Query coverage | E value |
|---|---|---|---|---|---|
| AEJ95084.1 | gp50 [Mycobacterium phage Gladiator] | 122 | 122 | 78% | 1e-33 |
| AEJ93380.1 | gp52 [Mycobacterium phage EricB] >gb\|AEK08494.1\| gp51 [Mycobac | 119 | 119 | 62% | 6e-33 |
| AEK08698.1 | gp53 [Mycobacterium phage Hammer] | 120 | 120 | 78% | 1e-32 |
| YP_655631.1 | gp52 [Mycobacterium phage Che12] >gb\|ABE67371.1\| gp52 [Mycoba | 109 | 109 | 61% | 4e-29 |

ORF2

**Figure 3.45** BLASTp search result of ORF 1 and 2 from clone 57 of Mozambique 3 *Pst*I library when aligned with known protein sequences from the database

Many of the predicted OFR did not have an assigned function, and as discussed above, many of the ORFs that have protein sequence similarity to those in the database are similar to those that are responsible for nucleotide production. This results, however, is also seen in another study on antimicrobial genes found in Rhodococcus phage YF1 (Shibayama and Dabbs 2011). In this study, 13 ORFs from 10 clones (with sizes ranging from 390bp to 960bp) were predicted with only three having functions assigned to them. One such ORF also have protein sequence similarity to those of thymidylate synthase. Although it has been suggested above that the presence of genes that partake in DNA replication helps in phage DNA proliferation, thereby increasing transcription of phage genes (including antimicrobial ones) and increasing the pace of infection, Shibayama and Dabbs has also argued that TS could have been encoded by YF1 genes with the purpose of upsetting the balance of enzymes in thymidine production, thus interfering with host DNA replication.

However, it is important to note that the enzymes discussed above are themselves antimicrobial targets for future drug development. As mentioned above, there are already several synthesized protein inhibitors for DNA primase and even one that is undergoing animal testing (Baumeister *et al* 2007), with TS being seen as another potential antimicrobial target (Matthews *et al* 2003).

Since the genomic library of various phages was built using *E. coli*, none of the clones in this study would be inhibitory to *E. coli*. Hence the question remains whether or not all phage clones in the genomic library represents all the genes in the phage as any such genes that are inhibitory to *E. coli* would be lost. However, it is not clear if any of the phage genes are only inhibitory to *Mycobacterium/Rhodococcus* or do these genes simply fail to express themselves in *E. coli* (Shbayama and Dabbs 2011). In any case, since these genomic libraries were constructed using various restriction endonuclease, genes that could exhibit antimicrobial activates but are located at the ends of the genome will not be included, as the ends of the genome will not have compatible ends to be incorporated into either pUC18/19 or pDA71.

**Table 3.16** Summary of screened clones and their predicted ORFs

| Clone # | ORF # | Start | End | Phage/Organism | Function | Protein size (aa) |
|---|---|---|---|---|---|---|
| M1 clone 5 | 1 | 1086 | 1496 | L5 gp27 | Minor tail subunit 38.5kd protein | 127 |
| | 2 | 1503 | 1805 | Trixie gp29 | Minor tail subunit | 100 |
| M1 clone 6 | 1 | 2 | 79 | Hammer gp37 | Hypothesised parA protein - chromosome segregation | 26 |
| | 2 | 184 | 351 | N/A | Unknown | 55 |
| M3 clone 22 | 1 | 1 | 310 | Redrock gp46 | DNA polymerase I - 3'-5' exonuclease and polymerase domain | 103 |
| | 2 | 319 | 483 | Hammer gp49 | Unknown | 54 |
| | 3 | 476 | 658 | Hammer gp49 | Unknown | 60 |
| | 4 | 1298 | 1596 | Benedict gp44 | Thymidylate synthase complementing protein (ThyX/Thy1) | 99 |
| M3 clone29 | 1 | 56 | 184 | N/A | Unknown | 42 |
| | 2 | 471 | 617 | Che12 gp63 | Endonuclease VII; pfam02945 | 48 |
| | 3 | 614 | 760 | Turbido gp61 | DNA primase zinc binding motif | 48 |
| M3 clone 42 | 1 | 492 | 674 | Turbido gp54 | Unknown | 60 |
| | 2 | 1125 | 1367 | Che 12 gp53 | Ribonucleotide reductase and pyruvate formate lyase | 80 |
| M3 clone 57 | 1 | 3 | 408 | Pukovnik gp48 | Unknown | 135 |
| | 2 | 474 | 758 | Gladiator gp50 | Unknown | 94 |

### 3.10 Yeast transformation

Clones were selected to test for inhibitory activities in eukaryotes. There have been a number of researches done on the nature of antimicrobial activity of certain phage genes. Some of the genes target the replication mechanism and also transcription and replication apparatus of the cell, details of which have been discussed in the introduction above. Known inhibitory clones were transformed into *S. cerevisiae* to observe if these inhibitory clones can be expressed in yeast and whether if inhibitory activity can also be observed in eukaryotes. Should the activities be observed, it could shed light on the nature of phage gene products. Such genes could be involved in the disruption of cell DNA or RNA metabolism common in both prokaryotes and eukaryotes.

Selected inserts from both *Pst*I libraries of Mozambique 1 and Mozambique 3 was digested with *Pst*I, ran on a 0.6% agarose gel, extracted from the gel and purified. These purified inserts were then ligated into pUC18/19 *Pst*I site and transformed into *E. coli* MM294-4. New clones of pUC18/19 were then collected by mini plasmid preparations and screened for the presence inserts. These inserts were then digested with enzymes *Hind*III and *Xba*I and extracted from pUC18/19 in the same manner as they were extracted from PDA71. Inserts are then ligated into yeast vector pSHLeu, transformed again into *E. coli* MM294-4 and clones were again screened for the presence of inserts. These inserts are shown below after being released from the yeast vector after double digestion with *Hind*III and *Xba*I

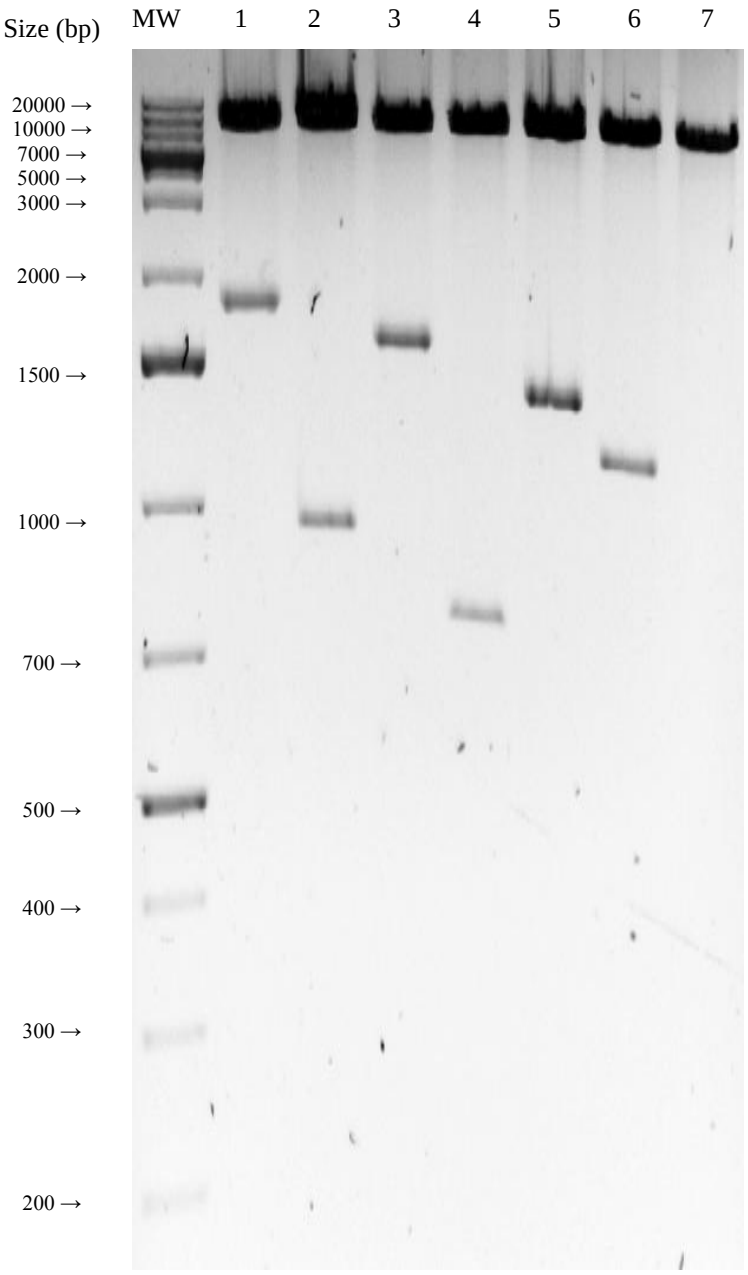
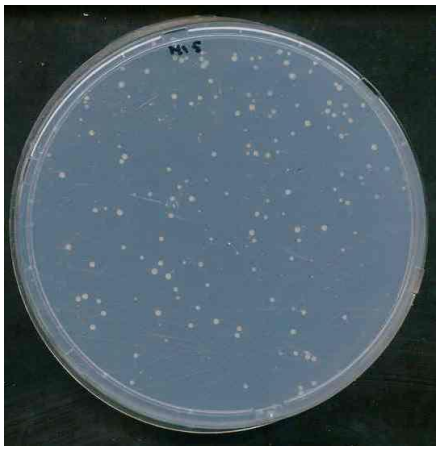
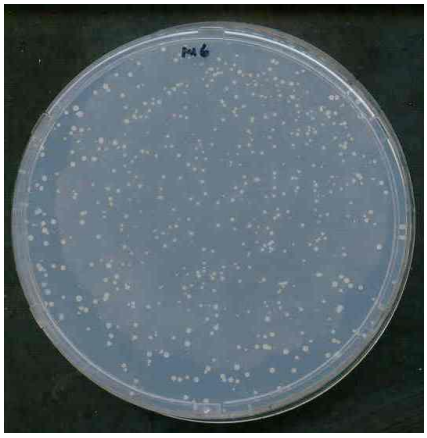



**Figure 3.46** Inserts of clones from Mozambique 1 and 3 *Pst*I libaries in yeast vector pSHLeu. Lane 1: clone 5 of *Pst*I library of Mozambique 1; lane 2: clone 6 of Mozambique 1; lane 3: clone 22 of *Pst*I library of Mozambique 3; lane 4: clone 29 of Mozambique 3; lane 5: clone 42; lane 6: clone 57 and lane 7: pSHLeu control

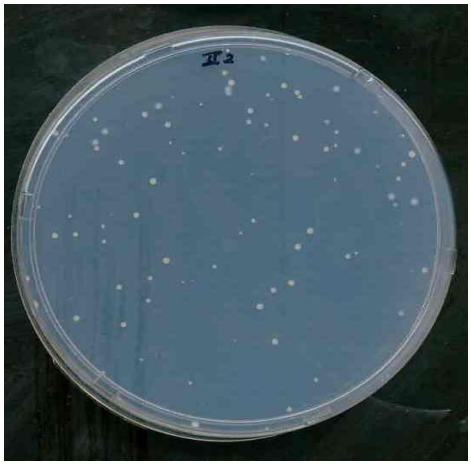
Inserts in vector pSHLeu was then transformed into *S. cerevisiae*



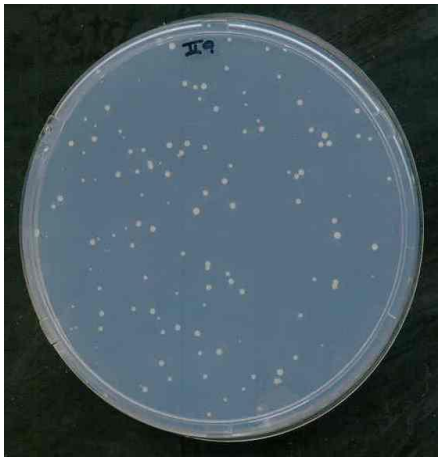
Clone 5 of Mozambique 1



Clone 6 of Mozambique 1

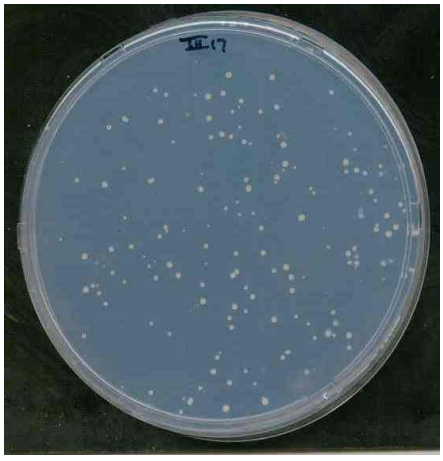


Clone 22 of Mozambique 3



Clone 29 of Mozambique3

Clone 42 of Mozambique 3



Clone 57 of Mozambique 3

pSHLeu control

**Fig 3.47** Transformation of *S. cerevisiae* by detected clones from *Pst*I libraries of Mozambique 1 and Mozambique 3 along with vector only control

**Table 3.17** Transformation efficiencies of yeast transformation for selected clones isolated from Mozambique 1 and Mozambique 3 *Pst*I libraries

| Clone # | Transformants/µg of DNA |
|---|---|
| 5 | 20 |
| 6 | 69 |
| 22 | 18 |
| 29 | 21 |
| 42 | 12 |
| 57 | 30 |
| pSHLeu Control | 308 |

Clones tested in *S. cerevisiae* yielded much less colonies than those of the control, with the lowest figure being 12 transformants/µl of DNA for clone 42 while those of the control yielded 308 transformants/µl of DNA. However, when this is compared with the inhibitory figures from those of *R. erythropolis*, the highest figure of inhibition is 12.5 transformants/µl of DNA, while its control yielded 1447 transformants/ µl of DNA. It is clear that when comparing the figures between the two transformations, the level of inhibition is observed to be much higher in *R. erythropolis* than in *S. cerevisiae*. Therefore, it could be argued that the clones do not exhibit inhibitory activities in *S. cerevisiae*, and that if it did, its effect is only partial and it is much less effective and much lower than those observed in *R. erythropolis*.

Reasons for the lack of inhibitory activity observed can be suggested based on the protein similarities of various ORFs to those proteins in the database with known functions or structures. In the case of ORF1 and 2 of clone 5, the protein sequences of the two ORFs are similar to those of phage tail subunits. And as mentioned above, are related to phage tail like bacteriocins that are produced by bacteria. There has been no record of such bacteriocins having an inhibitory effect on eukaryotes. Since these tail like peptides are descended from defunct phages and all phages are known to only infect their particular niche of bacterial hosts, it is to be expected that these antimicrobial proteins to have no effect on *S. cerevisiae*.

All other ORFs have protein similarity with proteins in the database that deals in DNA synthesis and metabolism, and unlike phage tail-like bacteriocins, such genes could be expected to display inhibitory activity towards *S. cerevisiae*. However, since the enzymes that are responsible for DNA synthesis are different between those found in eukaryotes and those in prokayotes, it is doubtful whether these ORFs, even if successfully expressed, will have much impact on *S. cerevisiae*. An example of this is the DNA primase: As discussed above, the structure of DNA primase found in T7 consists of one subunit with a single nucleotide binding site and two NTP binding sites (Kuchta and Stengel 2010) along with the ability to carry out helicase activity, while those in eukaryotes have two subunits and an iron-sulphur domain (Klinge *et al* 2007). In any

case, these phage genes have also failed to inhibit other prokaryotes, namely the *E. coli* host that was used to construct these genomic libraries.

More experiments, such as a repeat of the transformation of *S. cerecisiae*, or testing individual subclones could shed more light into exact nature of the clones and help determine more accurately the level of inhibition if there are any.

# 4. Appendices

## Appendix A – Media and Solutions

### 1 – Media

**LB**

1g tryptone
0.5g sodium chloride
0.5g yeast extract
100ml $dH_2O$

**LA**

1g tryptone
0.5g sodium chloride
0.5 yeast extract
1.5g agar
100ml $dH_2O$

**Sloppy agar**

1g tryptone
0.5g sodium chloride
0.5 yeast extract
0.75g agar
100ml $dH_2O$

**LBSG**

1g tryptone

0.5g sodium chloride

0.5g yeast extracts

10.3g sucrose

2g glycine

100ml $dH_2O$

***Rhodococcus* regeneration media**

3g tryptone

0.9g sodium chloride

1.5g yeast extract

30.9g sucrose

1g glucose

1g magnesium chloride hexahydrate

250ml $dH_2O$

The mixture was heated in a microwave to dissolve the sucrose. 5.5g of agar was then added, autoclaved and cooled to 60$^{o}$C and the following was added before being poured on a horizontal surface:

10ml 0.25 TES pH 7.2

6ml 1M calcium chloride

3ml 0.5% potassium dihydrogen orthophosphate

1.5 ml 10 mg/ml nystatin

0.9 ml 10 mg/ml rifampicin

**YPD**

2g peptone

1g yeast extract

2g glucose

**SC**

1.34g yeast nitrogen base (without amino acids)

4g glucose

0.088 dropout mix

0.012g L-histidine

0.012g L-tryptophan

0.004g uracil

**Dropout mix**

0.1g L-arginine

0.3g L-lysine

0.2g L-methionine

0.1g L-phenylalanine

0.3g L-serine

1.0g threonine

0.2g L-tyrosine

0.1g adenine

Mixture was dissolved in 5L of $dH_2O$

**1M Calcium chloride**

2.2g calcium chloride

10ml $dH_2O$

**1M Magnesium chloride**

2.03g magnesium chloride

10ml $dH_2O$

**2 – Solutions for phage purification**

**SC buffer**

0.58 g sodium chloride

0.2 g calcium chloride hexahydrate

5 ml 1 M Tris-hydrochloride pH 8.0

0.5 ml 2% gelatin

95 ml $dH_2O$

**2% Gelatin**

0.02 g gelatin

1 ml $dH_2O$

4.09g Cesium chloride was added to 5ml of SC buffer to achieve a density of 1.50g/ml.

**1 M Tris-hydrochloride pH 8.0**

12.1g Tris base dissolved in 50ml dH2O

HCl added to adjust pH to 8.0

$dH_2O$ added to a final volume of 100ml

**3 – Solutions for Transmission Electron Microscopy**

**0.35% Formvar**

0.35g Formvar

100ml chloroform

**1% Uranyl acetate**

0.05g uranyl acetate

5ml $dH_2O$

**4 – Solutions for phage DNA extraction**

**Dialysis buffer**

0.58g sodium chloride

2.2g calcium chloride hexahydrate

50ml 1M Tris-hydrochloride pH 8.0

950ml $dH_2O$

**5 – Solutions for DNA manipulations**

**Solution I**

0.9g glucose

2.5ml 1M Tris-hydrochloride pH 8.0

2ml 0.5M EDTA pH 8.0

95.5ml $dH_2O$

**Solution II**

0.8g sodium hydroxide

1g SDS

100ml $dH_2O$

**Solution III**

60ml 5M sodium acetate

11.5ml glacial acetic acid

28.5ml $dH_2O$

**0.5M EDTA pH 8.0**

18.61g EDTA dissolved in 50ml of dH2O

Sodium hydroxide added to adjust pH to 8.0

$dH_2O$ added to a final volume of 100ml

**TE buffer**

1ml 1M Tris-hydrochloride pH 8.0

2ml 0.5M EDTA pH 8.0

97ml $dH_2O$

**10% SDS**

1g SDS

10ml dH2O

**1M Sodium chloride**

0.58g Sodium chloride

10ml $dH_2O$

**1% Ribonuclease**

0.01g ribonuclease

1ml $dH_2O$

Boil for 10 min

**6 – Solutions for electrophoresis**

**5× TBE**

54g Tris base

27.5g boric acid

20ml 0.5M EDTA pH 8.0

$dH_2O$ added to a final volume of 1000ml

**Running buffer**

20ml 5× TBE

20µl 1% EtBr

200ml $dH_2O$

**Agarose gel**

0.32g (for 0.8%), 0.24g (for 0.6%) agarose

4ml 5× TBE

36ml $dH_2O$

**1% EtBr**

0.1g EtBr

10ml $dH_2O$

**Tracking Dye**

0.01g bromophenol blue

0.01g xylene cyanol

10ml 30% glycerol in TE buffer

**7 – Solutions for Transformation**

**Calcium transformation buffer**

4.38g calcium chloride hexahydrate

2ml 1M Tris-hydrochloride pH 8.0

198ml $dH_2O$

**Basal (B) buffer**

10.3g glucose

0.025g potassium sulphate

0.202g magnesium chloride hexahydrate

10ml 0.25M TES pH 7.2

87.5ml $dH_2O$

**Protoplast (P) buffer**

2ml B buffer

20µl 0.5% potassium dihydrogen orthophosphate

50µl 1M calcium chloride

**P-PEG buffer**

0.5g PEG 6000 UV sterilized for 10min

1ml P buffer

Vortexed vigorously to mix

**0.25 TES pH 7.2**

17.2g TES in 250ml $dH_2O$

Sodium hydroxide added to adjust pH to 7.2

$dH_2O$ added to a final volume of 300ml

**0.5% potassium dihydrogen orthophosphate**

0.5g potassium dihydrogen orthophosphate

100ml $dH_2O$

**1M Lithium acetate**

6.56g lithium acetate

**50% PEG4000**

0.5g PEG4000

1ml $dH_2O$

**8 – Antibiotics**

**Table 4.1** List of antibiotics used

| Antibiotic | Concentration (mg/ml) | Solvent |
|---|---|---|
| Ampicillin | 100 | 70% ethanol 30% dH2O |
| Chloramphenicol | 4 | ethanol |
| Kanamycin | 100 | dH2O |
| Nalidixic acid | 100 | 70% ethanol 30% dH2O |
| Nystatin | 10 | methanol |
| Rifampicin | 10 | methanol |
| Streptomycin | 100 | dH2O |

**Appendix B – Restriction Maps of Plasmids in this Study**

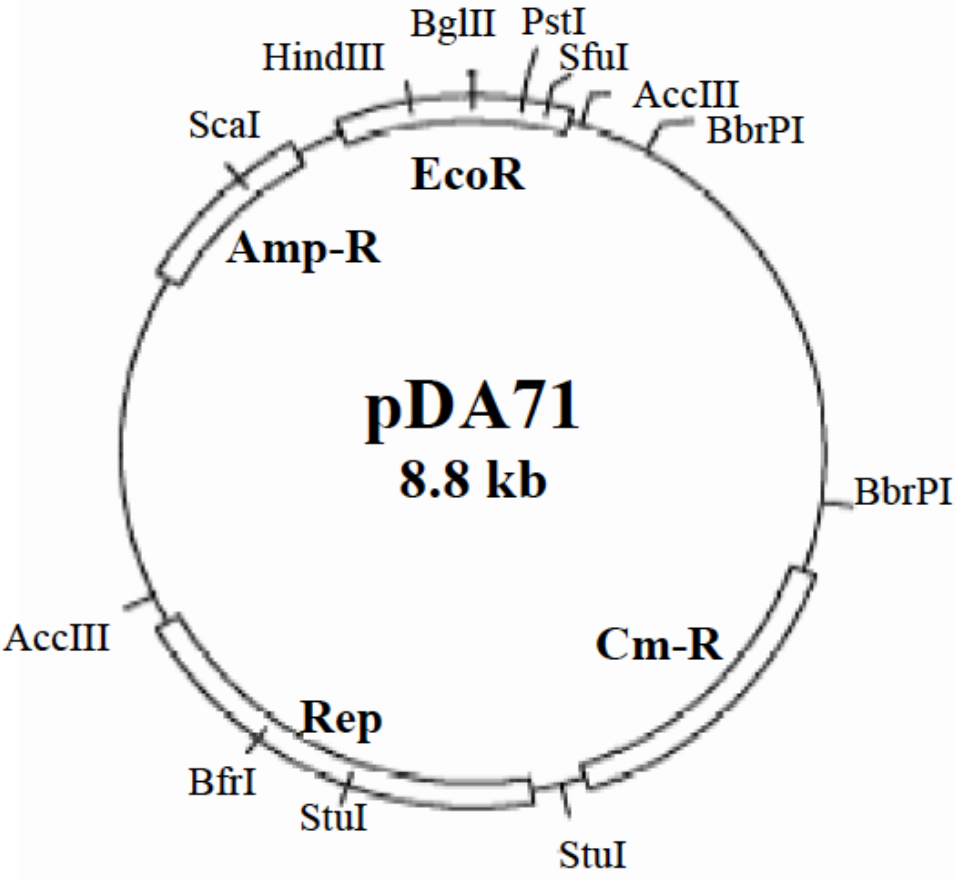



**Figure 4.1** Restriction map of pDA71

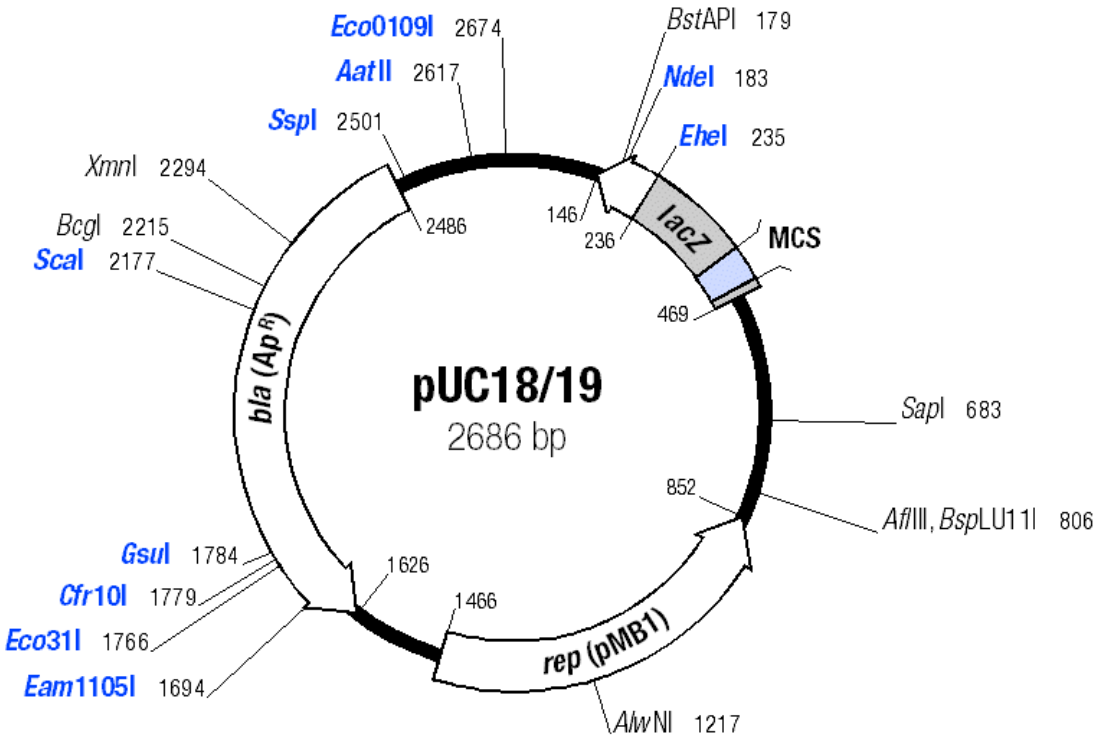
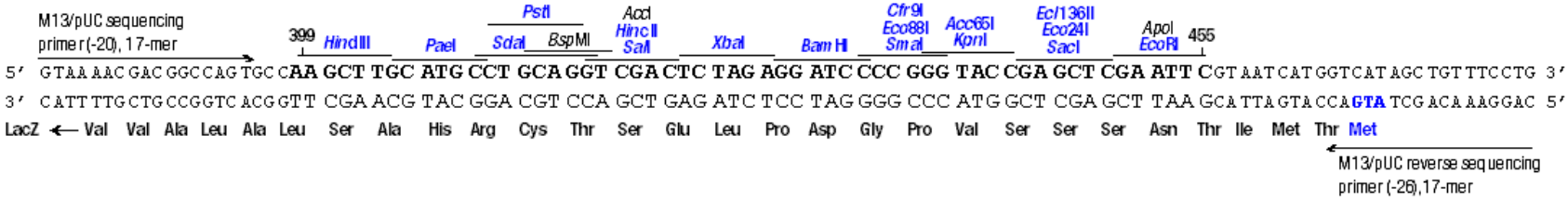


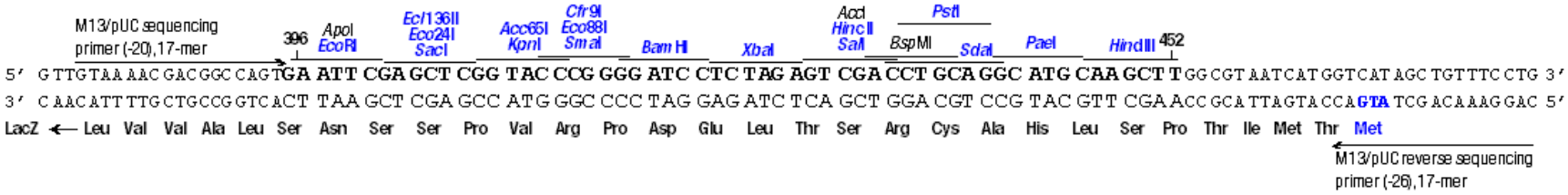

**Figure 4.2** Restriction map of pUC18/19

pUC18




1

pUC19




2

**Figure 4.3** Multiple cloning sites of 1) pUC 18 and 2) pUC 19

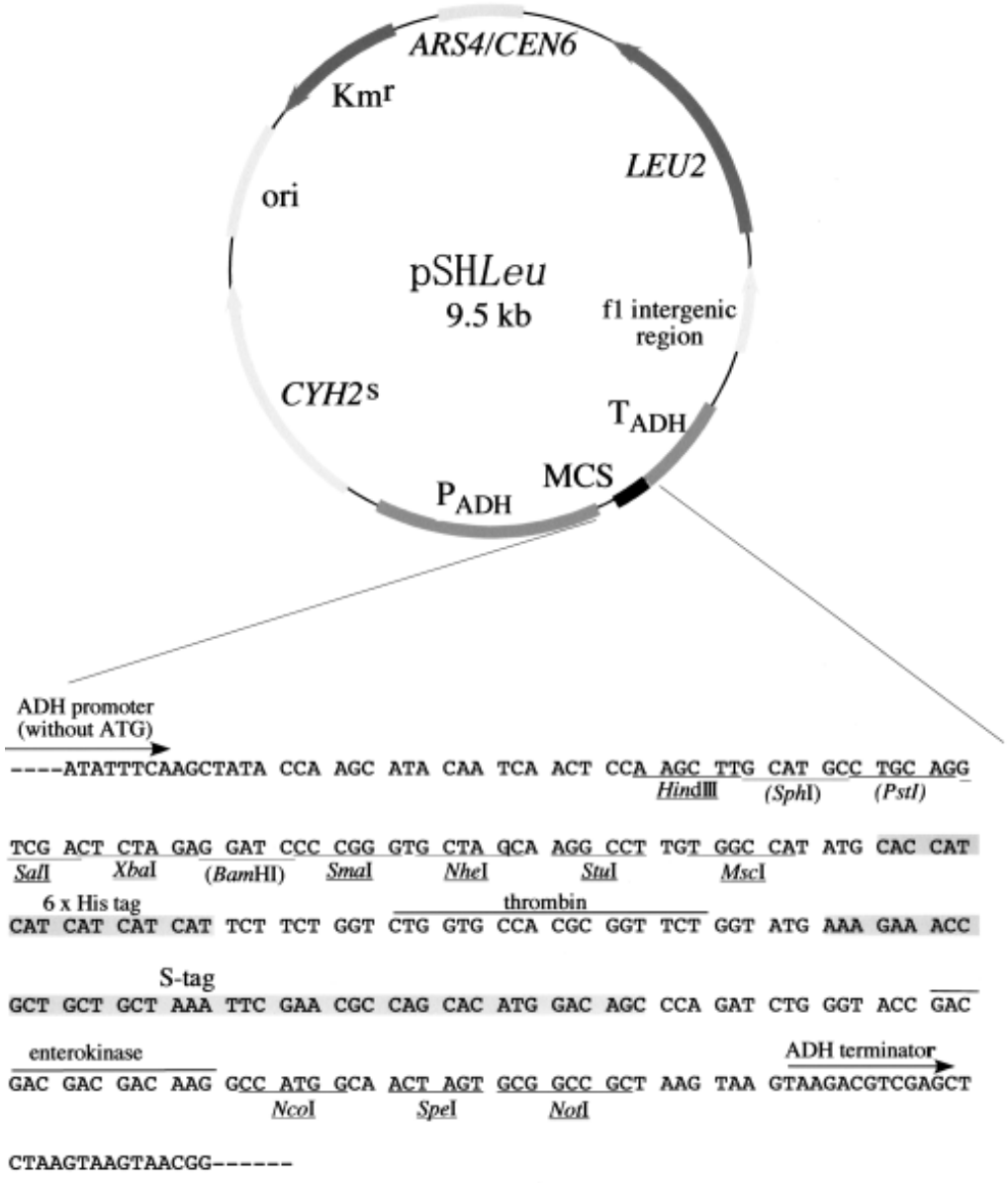
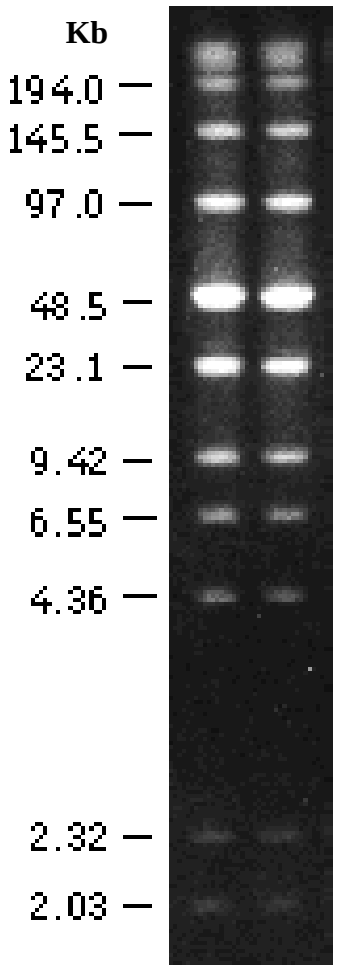
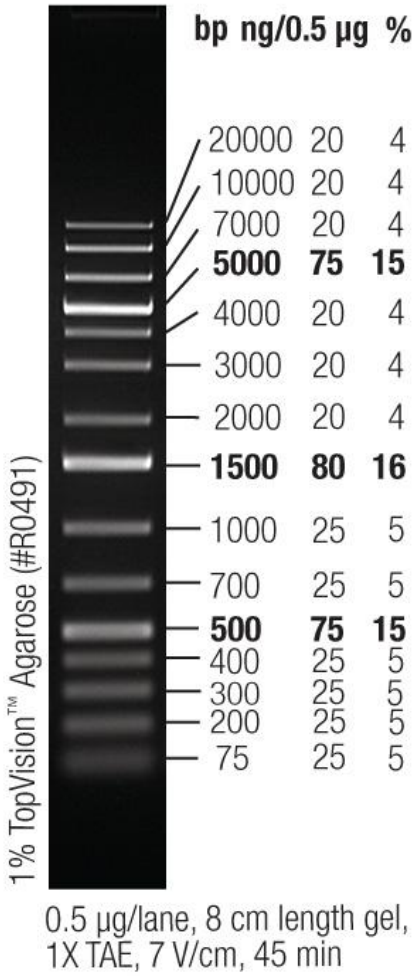



**Figure 4.4** Structure of pSHLeu and its multiple cloning sites. Unique restriction sites within the multiple cloning sites are underlined (adapted from Gan *et al* 2002)

**Appendix C Molecular Weight Markers**




1




2

**Figure 4.5** Molecular weight markers of 1) New England Biolabs pulse field gel electrophoresis low range marker and 2) Fermentas GeneRuler 1kB plus DNA ladder

## Appendix D – Properties of Proteins with Sequence Similarities with Phage Inibitory ORFs

**Table 4.2** Properties of L5 gp27 minor tail subunit

```
LOCUS       NP_039691                336 aa            linear   PHG 16-JUL-2008
DEFINITION  observed 38.5Kd protein [Mycobacterium phage L5].
ACCESSION   NP_039691
VERSION     NP_039691.1  GI:9625457
DBLINK      Project: 14459
DBSOURCE    REFSEQ: accession NC_001335.1
KEYWORDS    .
SOURCE      Mycobacterium phage L5
  ORGANISM  Mycobacterium phage L5
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
            L5-like viruses.
REFERENCE   1  (residues 1 to 336)
  AUTHORS   Donnelly-Wu,M.K., Jacobs,W.R. Jr. and Hatfull,G.F.
  TITLE     Superinfection immunity of mycobacteriophage L5: applications for
            genetic transformation of mycobacteria
  JOURNAL   Mol. Microbiol. 7 (3), 407-417 (1993)
   PUBMED   8459767

     Protein         1..336
                     /product="observed 38.5Kd protein"
                     /function="minor tail subunit"
                     /name="gene 27"
                     /calculated_mol_wt=38425

ORIGIN
        1 mitdtivele gvngerfnlt tgdqgvylat dvegcfydpp vkvvveepgn ypgarylshr
       61 alkrdivfgv vilndakqgp rswlsrdsew rkawafnrtc klyvttpdsg trylklalfe
      121 sptvkmdtdp rgkplevtvm sciaydpfwy eddkvfsakt ktdtrfdpsf wtppwpweel
      181 pketlrikvg reqgglnptd qyifpkwtvp gstekvpnfp wpfppnvpip wetapftqfv
      241 ipdysfedee frnrrlktpg liygencvid tdrreeqias esgspvwarm ngvrfrnsip
      301 pyteeaefvi dasgcapgqv vtlrltrpws rcwgle
```

**Table 4.3** Properties of Trixie gp29 minor tail subunit

```
LOCUS       AEL17859                 845 aa            linear   PHG 11-AUG-2011
DEFINITION  gp29 [Mycobacterium phage Trixie].
ACCESSION   AEL17859
VERSION     AEL17859.1  GI:342221265
DBSOURCE    accession JN408461.1
KEYWORDS    .
SOURCE      Mycobacterium phage Trixie
  ORGANISM  Mycobacterium phage Trixie
            Viruses; dsDNA viruses, no RNA stage; Caudovirales.
REFERENCE   1  (residues 1 to 845)
  AUTHORS   Ashfaq,M., Bobbity,S., Campbell,R., Chalissery,S., Cheekati,S.,
            Chitta,P.A., Hayes,C., Johnson,L., Karuturi,S., Mistry,N., Oh,S.,
            Patil,P., Reddy,R., Saxena,R., Siegfried,S., Wang,H.-Y., Ward,L.,
            Williamson,S.M., Anderson,J., Padolina,J., Johnson,A.,
            Serrano,M.G., Lee,V., Hendricks,S.L., Sheth,N.U., Buck,G.A.,
            Bradley,K.W., Khaja,R., Lewis,M.F., Barker,L.P., Jordan,T.C.,
            Russell,D.A., Pope,W.H., Jacobs-Sera,D., Hendrix,R.W. and
            Hatfull,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JUL-2011) Pittsburgh Bacteriophage Institute and
            Department of Biological Sciences, University of Pittsburgh, 365
            Crawford Hall, Pittsburgh, PA 15260, USA
     Protein         1..845
                     /product="gp29"
                     /name="tapemeasure"

ORIGIN
        1 mpnsagveva risvkvspnt kefrrelkad lekiekemsa dievhadlha aqtkadfrrl
       61 mtelkteaak gvevdvhvnk kggilgglfg gkggkspsse igdigdeaek ttqkvlsmgq
      121 gflgmsrmaw igvgvialaa pavalvagll aglpslmaaf gagaavvalg mdgikaaast
      181 iapvldqvkt qvsavfkegl tpvmgqlgtm ltsitpglkn vagslvamag gitdvvskgv
      241 gldqinnilg ktgdffrglt pvistgvqsf ltlanagans fntllaplqt fatqfndmvn
      301 ritsngqfga amqglssvlg svlnlftrlf esgvqamgql ggplstling fgdafialmp
      361 altslsslla nvlgtalsal apaitaltpa ftmladtlgt llvsniqsla piltqvatli
      421 ggtlttalqa iqpmlpglvq sfaqlsqtlv tqlapyipql atafgqiaga vlqlvptlvs
      481 alipafqqml pailqlvpsl vsmvqsfari mpvvvpvvei iiklaaaviq agasiasfli
      541 gglsrlvgil aevtsavttw vsswsngvqq vsdyvgqlpg kikswfddag swlieagkni
      601 vqglingigs missavskak elasgvknav tsflgihsps kvfeqlgiyt gqgyaigldk
      661 gfapvleqak alasqitaav asgtedptal lngfskqdvs rmekvlgtei kkyerqakal
      721 dlqakatgdd slkaeaqklr dmkdqlqtqk emldlagdyn detssgngag slesqvsklm
      781 aspvdfakat gkqflsdigi sgegflskli tegtqyifqi gsvdealsik dreesksals
      841 vvgrq
```

**Table 4.4** Properties of Hammer gp37 parA-like protein

```
LOCUS       AEK08682                228 aa            linear   PHG 17-JUL-2011
DEFINITION  gp37 [Mycobacterium phage Hammer].
ACCESSION   AEK08682
VERSION     AEK08682.1  GI:339782963
DBSOURCE    accession JF937094.1
KEYWORDS    .
SOURCE      Mycobacterium phage Hammer
  ORGANISM  Mycobacterium phage Hammer
            Viruses; unclassified phages.
REFERENCE   1  (residues 1 to 228)
  AUTHORS   Luchansky,M.S., Bowman,C.A., Hinkle,K.M., Howard,J.T., Lawson,C.W.,
            Mauer,R.M., Rosenfield,N.R., Mattson,V., Ko,C., Russell,D.A.,
            Pope,W.H., Jacobs-Sera,D., Hendrix,R.W. and Hatfull,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAY-2011) Pittsburgh Bacteriophage Institute and
            Department of Biological Sciences, University of Pittsburgh, 365
            Crawford Hall, Pittsburgh, PA 15260, USA
COMMENT     Method: conceptual translation.
FEATURES             Location/Qualifiers
     source          1..228
                     /organism="Mycobacterium phage Hammer"
                     /strain="Hammer"
                     /isolation_source="soil"
                     /db_xref="taxon:1034103"
                     /lab_host="Mycobacterium smegmatis mc2 155"
                     /country="USA: Pittsburgh, PA"
                     /collection_date="21-Jun-2005"
                     /collected_by="M. Luchansky"
                     /identified_by="M. Luchansky"
     Protein         1..228
                     /product="gp37"
                     /name="ParA"
ORIGIN
        1 msscqvkspr vlessslvl gvssprlfee ssnltsmtti svvhtkggvg ktttamylat
       61 aaarsgvdtv vvdadpqksa skwvntaaer gigmpfevvd gsdrlvlpdk elvlvdtppg
      121 tsdliqqaia gadlviipcg aspievervf ptlnltvqvp slvlmtqvdl rakmvdrvrg
      181 mfktrgvpvf ntmvthrqsi kksfgtmpvd leaywdvwge legvvvgv
```

**Table 4.5** Properties of Redrock gp46family A polymerase

```
LOCUS       ADB93739                 595 aa            linear   PHG 01-FEB-2010
DEFINITION  gp46 [Mycobacterium phage RedRock].
ACCESSION   ADB93739
VERSION     ADB93739.1  GI:284794813
DBSOURCE    accession GU339467.1
KEYWORDS    .
SOURCE      Mycobacterium phage RedRock
  ORGANISM  Mycobacterium phage RedRock
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
REFERENCE   1  (residues 1 to 595)
  AUTHORS   Jacobs-Sera,D., Zellars,M., Wells,M.E., Webb,J.L., Ware,V.C.,
            Vazquez,E., TamarapuParthasarathy,P., Smith,I.A., Simon,S.E.,
            Shaffer,C.D., Rubin,M.R., Rosenzweig,R.F., Rinehart,C.A., Qin,H.,
            Pillay,I., Payne,D.E. II, Padolina,J.M., Novick,P.A., Miller,E.S.,
            Mayer,E.S., Marzillier,J.Y., Mageeney,C.M., MacGibeny,M.A., Li,W.,
            Lee,J.Y., Kinnersley,M.A., King-Smith,C., King,R.A., Kenna,M.A.,
            Kearse,M.G., Johnson,B.K., Johnson,A.A., Johnson,C.M., Hughes,L.E.,
            Harrison,M., Guild,N.A., Gilbert,J.L., Fillman,C.L., Felton,C.M.,
            Dunbar,D.A., Dennehy,J.J., DeJong,R.J., Carson,S., Burnett,S.H.,
            Breakwell,D.P., Berrios,J.E., Benjamin,R.C., Anderson,J.J.,
            Bradley,K.W., Khaja,R., Lee,E., Barker,L.P., Lewis,M.F.,
            Jordan,T.C., Cresawn,S.G., Grace,M.A., Pope,W.H., Ko,C.,
            Russell,D.A., Peebles,C.L., Lawrence,J.L., Hendrix,R.W. and
            Hatfull,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-DEC-2009) Department of Biological Sciences and
            Pittsburgh Bacteriophage Institute, University of Pittsburgh, 365
            Crawford Hall, Pittsburgh, PA 15260, USA
     Protein         1..595
                     /product="gp46"
     Region          19..571
                     /region_name="PolA"
                     /note="DNA polymerase I - 3'-5' exonuclease and polymerase
                     domains [DNA replication, recombination, and repair];
                     COG0749"
                     /db_xref="CDD:31092"
ORIGIN
        1 mielrhevqg dlvtinvvet pedlagfrdf irahlnclav dtettgldiy sdtfecrlvq
       61 fgtqdeawvv pvelgdvfie dvriaigalr kvvlqnasyd lqvldqcfgi emedlwprvl
      121 dtqilaklvd prpfeaggfg hsleeliaef iskeqaetvk klmaklaqeh kttkakiwat
      181 idlfhpeylk yagmdtifta rvckaltplv pdvsrglvay eheiseicsy idrqgflldv
      241 eyaqelsdkw ladqqvweai afteygvekv nstedlaegl eemgvkitgr tetgkrqvdk
      301 tlleklvaeg nelatiaqea kklgkwrktw vqkfldtrda edrchtfvnp lqartsrmsi
      361 tgipaqtlpa sdstvrrcfl aeegdvmasi dyqtqelrvl aalsndqtmi eafktgadlh
      421 qmtadaaqvt rkvgkmanfl tvygggaktl aeqakvdfpt akrtldafar typgvarlsk
      481 klgteagrkg yivtpvgrrl pvdssrsysa lnymvqsssr dvtcralirl hkagytpylr
      541 lpihdeivas lpaseaeraa ahigqlmaeq mgpvligtdp evgkrswgsl ygady
```

**Table 4.6** Properties of Benedict gp44 thymidylate synthase

```
LOCUS       AEJ93473                235 aa            linear   PHG 11-AUG-2011
DEFINITION  gp44 [Mycobacterium phage Benedict].
ACCESSION   AEJ93473
VERSION     AEJ93473.1  GI:339753449
DBSOURCE    accession JN083852.1
KEYWORDS    .
SOURCE      Mycobacterium phage Benedict
  ORGANISM  Mycobacterium phage Benedict
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
REFERENCE   1  (residues 1 to 235)
  AUTHORS   Alvarez,F.A., Cornejo,C.M., Damoiseaux,J.S., Dossey,J., Drum,B.E.,
            Fashina,O., Greenstein,S.K., Howard,C.M., Johnson,R.E., Kang,K.E.,
            Kuhn,J.L., Meese,A.M., Nelson,V.A., Nguyen,K., Orr,A.M.,
            Raley,M.A., Rosenthal,E.L., Rosenthal,M.K., Smith,L.D., Sotelo,J.,
            White,L., Wilson,A.C., Yanney,R.D., Wu,H., Adair,T.L., Gibbon,B.C.,
            Lee,V., Hendricks,S.L., Voegtly,L.J., Wang,Y., Glascock,A.L.,
            Anderson,J., Williamson,S.M., Walstead,R.N., Carvalho,M.R.C.,
            Johnson,A., Buck,G.A., Bradley,K.W., Khaja,R., Lewis,M.F.,
            Barker,L.P., Jordan,T.C., Russell,D.A., Pope,W.H., Jacobs-Sera,D.,
            Hendrix,R.W. and Hatfull,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2011) Pittsburgh Bacteriophage Institute and
            Department of Biological Sciences, University of Pittsburgh, 365
            Crawford Hall, Pittsburgh, PA 15260, USA
     Protein          1..235
                      /product="gp44"
                      /name="ThyX"
ORIGIN
        1 mkvellantt inpdalgtlr nlgyetrwds sasdelaefa grncyrsfdr pnpatrenvd
       61 ylkhilevgh esvlehasat fyieasrsvl telerhrhls fsvvsqryvd ptslgvhypp
      121 amnltegtsa hqkvldaidd vrflaldryk tivkelsalg lprkqareaa ravlpnmtns
      181 pmvvtgnhra wrnvikqrwh taadaeirel agellkqlrq vapntyqdip eepyg
```

**Table 4.7** Properties of Che12 gp63endonuclease

```
LOCUS       YP_655642                151 aa            linear   PHG 19-DEC-2007
DEFINITION  gp63 [Mycobacterium phage Che12].
ACCESSION   YP_655642
VERSION     YP_655642.1  GI:109392513
DBLINK      Project: 17143
DBSOURCE    REFSEQ: accession NC_008203.1
KEYWORDS    .
SOURCE      Mycobacterium phage Che12
  ORGANISM  Mycobacterium phage Che12
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
REFERENCE   1  (residues 1 to 151)
  AUTHORS   Hatfull,G.F., Pedulla,M.L., Jacobs-Sera,D., Cichon,P.M., Foley,A.,
            Ford,M.E., Gonda,R.M., Houtz,J.M., Hryckowian,A.J., Kelchner,V.A.,
            Namburi,S., Pajcini,K.V., Popovich,M.G., Schleicher,D.T.,
            Simanek,B.Z., Smith,A.L., Zdanowicz,G.M., Kumar,V., Peebles,C.L.,
            Jacobs,W.R. Jr., Lawrence,J.G. and Hendrix,R.W.
  TITLE     Exploring the mycobacteriophage metaproteome: phage genomics as an
            educational platform
  JOURNAL   PLoS Genet. 2 (6), E92 (2006)
   PUBMED   16789831

     Protein          1..151
                      /product="gp63"
                      /calculated_mol_wt=17185
     Region           48..129
                      /region_name="Endonuclease_7"
                      /note="Recombination endonuclease VII; pfam02945"
                      /db_xref="CDD:111794"
ORIGIN
        1 makarkprfc vdclaegitt krdakypgpr cythhrakrr erssgawgtr ilatygitpe
       61 eywkiyefqg grcyicqran gkkkrlsvdh dhktgivrgl lctmcnkytl gwardcieff
      121 kraiayltnp pavqvigeri apieadklsq t
```

**Table 4.8** Properties of Turbido gp61DNA primase

```
LOCUS       AEL17806                 133 aa            linear   PHG 11-AUG-2011
DEFINITION  gp61 [Mycobacterium phage Turbido].
ACCESSION   AEL17806
VERSION     AEL17806.1  GI:342221211
DBSOURCE    accession JN408460.1
KEYWORDS    .
SOURCE      Mycobacterium phage Turbido
  ORGANISM  Mycobacterium phage Turbido
            Viruses; dsDNA viruses, no RNA stage; Caudovirales.
REFERENCE   1  (residues 1 to 133)
  AUTHORS   Frey,S., Wilder,L., Trepanier,M., Childs,M., Brennan,C.,
            Messina,E., Kinnersley,M., Rosenzweig,F., Gates,C., Holben,W.E.,
            Gilbert,J., Trauscht,D., Flood,M.D., Anderson,J., Johnson,A.,
            Padolina,J., Serrano,M.G., Lee,V., Hendricks,S.L., Sheth,N.U.,
            Buck,G.A., Bradley,K.W., Khaja,R., Lewis,M.F., Barker,L.P.,
            Jordan,T.C., Russell,D.A., Pope,W.H., Jacobs-Sera,D., Hendrix,R.W.
            and Hatfull,G.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-2011) Pittsburgh Bacteriophage Institute and
            Department of Biological Sciences, University of Pittsburgh, 365
            Crawford Hall, Pittsburgh, PA 15260, USA

     Protein          1..133
                      /product="gp61"
                      /name="DNA primase"
     Region           31..85
                      /region_name="zf-CHC2"
                      /note="CHC2 zinc finger; cl15369"
                      /db_xref="CDD:199505"
ORIGIN
        1 mstpscsspi aqviqryypd weappdhnew nkclcpfhgd ekpsaavsyd lqgfnclacg
       61 vrgdvisiir heeevsfaea qriaadlsvg ghnqvprkpe rkssrrvfge srpartprst
      121 vrtgirgrpt pws
```

**Table 4.9** Properties of Che12 gp53Ribonucleotide reductase and pyruvate formate

```
LOCUS       YP_655632                672 aa            linear   PHG 19-DEC-2007
DEFINITION  gp53 [Mycobacterium phage Che12].
ACCESSION   YP_655632
VERSION     YP_655632.1  GI:109392503
DBLINK      Project: 17143
DBSOURCE    REFSEQ: accession NC_008203.1
KEYWORDS    .
SOURCE      Mycobacterium phage Che12
  ORGANISM  Mycobacterium phage Che12
            Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
REFERENCE   1  (residues 1 to 672)
  AUTHORS   Hatfull,G.F., Pedulla,M.L., Jacobs-Sera,D., Cichon,P.M., Foley,A.,
            Ford,M.E., Gonda,R.M., Houtz,J.M., Hryckowian,A.J., Kelchner,V.A.,
            Namburi,S., Pajcini,K.V., Popovich,M.G., Schleicher,D.T.,
            Simanek,B.Z., Smith,A.L., Zdanowicz,G.M., Kumar,V., Peebles,C.L.,
            Jacobs,W.R. Jr., Lawrence,J.G. and Hendrix,R.W.
  TITLE     Exploring the mycobacteriophage metaproteome: phage genomics as an
            educational platform
  JOURNAL   PLoS Genet. 2 (6), E92 (2006)
   PUBMED   16789831

     Protein          1..672
                      /product="gp53"
                      /calculated_mol_wt=74733
     Region           11..671
                      /region_name="RNR_PFL"
                      /note="Ribonucleotide reductase and Pyruvate formate
                      lyase; cl09939"
                      /db_xref="CDD:186877"

ORIGIN
        1 mtaeipwgpt gelvynrtya rtkpdgskes wpetvervvd gnlslvdery revderehlv
       61 rlitefkllp agrhlwasgv knaqhlfncw vsgwpekisd hfqftfmrlm egggvganys
      121 nhylegypav vhplkveivc dpdhvdyqam kdagilsehy dsdwvgafpi edsregwaaa
      181 mvdlidthyr ydtvhfqrvy dvsrirpqga klktfggvas gplpfaqmlq kvsevlsdra
      241 gevltgldam avdhaiaqcv vaggvrrsar msmmhwadpq itefinckse sgdhwttnis
      301 vevddqfwre vkspsestah dvmdalskga vrngepgmwd sslsnvgepn rvvctnpcge
      361 itlepwepcn lghinlaafv tpsgktdyid lirahrlmtr fliratfspv adpksrevld
      421 rnrrigvghl gvasylaltg rryseapgdk rftnflrema aevdlaaeqf shelripvpv
      481 kkrtiaptgt vakmpgvseg ihpifskyfi rrirfnensd maelrklane gyeiekdlfa
      541 pntevvsipt kdtlvqavvd lhgrdaeeiv eaandltlhq llafqalyqt cwadnavsft
      601 anvdpdayea vdvaadlkrf aglikgstif peasfpqapy eritkqqyes avakavsdgv
      661 deecangacp ik
```

## 5. References

Ackermann H.W. 2007. 5500 Phages Examined in the Electron Microscope. Archives of Virology Vol. 152 Pg. 227 – 243

Bair C.L, Rifat, D and Black L.W. 2007. Exclusion of Glucosyl-Hydroxymethylcytosine DNA Containing Bacteriophages is Overcome by the Injected Protein Inhibitor IPI*. Journal of Molecular Biology Vol. 366 Pg. 779 – 789

Baumeister J, Fischer R, Eckenberg P, Henninger K, Rubsamen-Waigmann H and Kleymann G. 2007. Superior Efficacy of Helicase–Primase Inhibitor BAY57-1293 for Herpes Infection and Latency in the Guinea Pig Model of Human Genital Herpes Disease. Antiviral Chemistry and Chemotherapy Vol.18 Pg. 35–48

Bergh O, Børsheim K.Y, Bartbak G and Heldal M. 1989. High Abundance of Viruses Founf in Aquatic Environments. Nature Vol. 340 Pg. 467 – 468

Bickle T.A and Krüger D.H. 1993. Biology of DNA Restriction. Microbiology Reviews Vol. 57 Pg. 434 – 450

Blanco L and Salas M. 1996. Relating Structure to Function in Φ29 DNA Polymerase. Journal of Biological Chemistry Vol. 271 Pg. 8509 - 8512

Boyer H.W. 1971. DNA Restriction and Modification Mechanisms in Bacteria. Annual Review of Microbiology Vol. 25 Pg. 153 – 176

Braibant M and Gilot P, Content J. 2000. The ATP Binding Cassette (ABC) Transport Systems of *Mycobacterium tuberculosis*. FEMS Microbiology Reviews Vol. 24 Pg 449 – 467

Broudy T.B and Fischetti V.A. 2003. In Vivo Lysogenic Conversion of Tox(-) *Streptococcus pyogenes* to Tox(+) with Lysogenic Streptococci or Free Phages. Infection and Immunity Vol. 71 Pg. 3782 – 3786

Breitbart M, Salamon P, Andresen B, Mahaffy J.M, Segall A.M, Mead D, Azam F and Rohwer F. 2002. Genomic Analysis of Uncultured Marine Viral Communities. Proceedings of the National Academy of Science of the United States of America Vol. 99 Pg. 14250 – 14255

Breitbart M, Felts B, Kelley S, Mahaffy J.M, Nulton J, Salamon P and Rohwer F. 2004. Diversity and Population Structure of a Near-Shore Marine-Sediment Viral Community. Proceedings Biological Sciences Vol. 271 Pg. 565 – 574

Brennan P.J and Nikaido H. 1995. The Envelope of Mycobacteria Annual Review of Biochemistry Vol. 64 Pg. 29 – 63

Brown E. 2004. Drugs against Superbugs: Private Lessons from Bacteriophages. TRENDS in Biotechnology Vol. 22 No. 9 Pg. 434 – 436

Carlson K, Raleigh E.A and Hattman S. 1994. Restriction and Modification. Molecular Biology of Bacteriophage T4. Karam J.D. (ed.). Washington DC. USA: American Society for Microbiology Press, Pg. 369 – 381

Casjens S. 2003. Prophages in Bacterial Genomics: What Have We Learnt So Far? Molecular Microbiology Vol. 49 Pg. 277 – 300

Casjens S.R. 2005. Comparative Genomics and evolution of the Tailed-Bacteriphages. Current Opinion in Microbiology Vol. 8 Pg. 451 – 458

Chambers H.F, Moreau D, Yajko D, Miick C, Wagner C, Hackbarth C, Kocagöz, Rosenberg E, Hadley W.K and Nikaido H. 1995. Can Penicillins and Other β-lactam Antibiotics be used to Treat *Tuberculosis*? Antimicrobial Agents and Chemotherapy Vol. 39 Pg. 2620 – 2624

Chanishvilli N. 2009. A Literature Review of Practical Application of Bacteriophages Research. Eliava Institute of Bacteriophages, Microbiology and Virology

Chen J, Kriakov J, Singh A, Jacobs W.R. Jr, Besra G.S and Bhatt A. 2009. Defects in Glycopeptidolipid Biosynthesis Confer Phage I3 Resistance in *Mycobacterium smegmatis*. Microbiology Vol. 155 Pg. 4050 – 57

Davis B.M and Waldor M.K. 2003. Filamentous Phages Linked to Virulence of *Vibrio cholerae*. Current Opinion in Microbiology Vol. 6 Pg. 35 – 42

De Rossi E, Arrigo P, Bellinzoni M, Silva P.A.E, Martin C, Ainsa J.A, Guglierame P and Riccardi G. 2002. The Multidrug Transporters Belonging to Major Facilitator Superfamily in *Mycobacterium tuberculosis*. Molecular Medicine Vol. 8 Pg. 714 – 724

Dye C and Williams B.G. 2010. The Population Dynamics and Control of *Tuberculosis*. Science Vol. 328 Pg. 856 – 861

Easter J and Gober J.W. 2002. ParB-Stimulated Nucleotide Exchange Regulates a Switch in Functionally Distinct ParA Activities. Molecular Cell Vol.10 Pg. 427 – 434

Eklund H, Uhlin U, Färnegårdh M, Logan D.T and Nordlund P. 2001. Structure and Function of the Radical Enzyme Ribonucleotide Reductase. Progress in Biophysics and Molecular Biology Vol. 77 Pg.177 – 268

Fischetti V.A. 2005. Bacteriophage Lytic Enzymes: Novel Anti-Infectives. TRENDS in Microbiology Vol. 13 Pg. 491 – 496

Fleischmann R.D, Adams M.D, White O, Clayton R.A, Kirkness E.F, Kerlavage A.R, Bult C.J, Tomb J.F, Dougherty B.A, Merrick J.M, McKenny K, Sutton G, FitzHugh W, Fields C, Gocyne J.D. Scott J, Shirley R, Liu L, Glodek A, Kelly J.M, Weidman J.F,

Flores A.R, Parson L.M, and Pavelka M.S. Jr. 2005. Genetic Analysis of the β-lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and Susceptibility to β-lactam Antibiotics. Microbiology Vol. 151 Pg. 521 - 532

Frederickson C.M, Short S.M and Suttle C.A. 2003. The Physical Environment Affects Cyanophage Communities in British Columbia Inlets. Microbial Ecology Vol. 46 Pg 348 – 357

Frick D.N and Richardson C.C. 2001. DNA Primases. Annual Review of Biochemistry. Vol. 70 Pg. 39 – 80

Friedman D.I, Olson E.R, Georgopoulos C, Tilly K, Herskowitz I, and Banuett F. 1984 Interactions of Bacteriophage and Host Macromolecules in the Growth of Bacteriophage Lambda. Microbiology Review Vol. 48 Pg 299 – 325.

Gan X, Kitakawa M, Yoshino K, Oshiro N, Yonezawa K and Isono K. 2002. Tag-Mediated Isolation of Yeast Mitochondrial Ribosome and Mass Spectrometric Identification of its New Components. European Journal of Biochemistry Vol. 269 Pg. 5203 – 5214

Glickman M.S and Jacobs W.R. Jr. 2001. Microbial Pathogenesis of *Mycobacterium tuberculosis*: Dawn of a Discipline. Cell. Vol. 104 Pg. 477 – 485

Gordon G.S and Wright A. 2000. DNA Segregation in Bacteria. Annual Review in Microbiology Vol. 54 Pg. 681 – 708.

Gratia J-P and Macrocin G: A Phage Tail Related to, but Distinct from the Temperate Coliphage Phi-Gamma and Produced by *Serratia marcescens* SMG 38. 1987. Archive of International Physiology and Biochimenistry Vol. 95: B30

Graziani S, Xia Y, Gurnon J.R, Van Etten J.L, Leduc D, Skouloubris S, Myllykallio H and Liebl U. 2004. Functional Analysis of FAD-Dependent Thymidylate Synthase ThyX from *Paramecium bursaria* Chlorella Virus-1. Journal of Biological Chemistry. Vol. 279 Pg. 54340 – 54347

Grebenshchikova S.M.L, Plugina L.A and Shcherbakov V.P 1994. The Role of T4-Bacteriophage Endonuclease VII in Correction of Mismatched Regions. Genetika Vol. 30 Pg. 622 – 626.

Hatfull G.F. 2010. Mycobacteriophages Genes and Genomes. Annual Review in Microbiology Vol. 64 Pg. 331 – 356

Hatfull G.F, Jacobs-Sera D, Lawrence J.G, Pope W.H. Russell D.A, Ko C.C, Weber R.J, Patel M.C, Germane K. L, Edgar R.H, Hoyte N.N, Bowman C.A, Tantoco A.T, Paladin E.C, Myers M.S, Smith A.L, Grace M.S, Pham T.T, O'Brien M.B, Vogelsberger A.M, Hryckowian A.J, Wynalek J.L, Donis-Keller H, Bogel M.W, Peebles C.L, Cresawn S.G and Hendrix R.W. 2010. Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition and Gene Size. Journal of Molecular Biology Vol. 397 Pg. 119 – 143

Hegde S.S, Vetting M.W, Roderick S.L, Mitchenall L.A, Maxwell A, Takiff H.E and Blanchard J.S. 2005. A Fluoroquinolone Resistance Protein from *Mycobacterium tuberculosis* that Mimics DNA. Science Vol. 308 Pg. 1480 – 1483

Hendrix R.W. 2003. Bacteriophage Genomics. Current Opinion in Microbiology Vol. 6 Pg. 506 – 511

Hendrix R and Casjens S. 2005. Bacteriophages λ and its Genetic Neighborhood. The Bacteriophages. Edn . Calender R (Ed.). Oxford Press.

Jarlier V and Nikaido H. 1994. Mycobacterial Cell Wall: Structural and Role in Natural Resistance to Antibiotics. FEMS Microbiology Letters Vol. 123 Pg. 11 – 18

Jordan A and Reichard P. 1998. Ribonucleotide Reductases. Annual Review of Biochemistry Vol. 67 Pg. 71– 98

Kemper B. 1997. Branched DNA Resolving Enzymes. In Nickoloff J.A.and Hoekstra M.F (eds), DNA Damage and Repair. Vol.1 DNA Repair in Prokaryotes and Lower Eukaryotes Pg. 179–204

Klinge S, Hirst J, Maman J.D, Krude T and Pellegrini L. 2007. An Iron–Sulfur Domain of the Eukaryotic Primase is Essential for RNA Primer Synthesis, Nature Structure and Molecular Biology Vol. 14 Pg. 875–877

Kolberg M, Strand K.R, Graaf P and Andersson K.K. 2004. Structure, function and Mechanism of Ribonucleotide Reductases. Biochimica et Biophysica Acta Vol. 1699 Pg. 1 – 34

Koul A, Arnoult E, Lounis N, Guillemont J and Andries K. 2011. The Challenge of New Drug Discovery for *Turberculosis*. Nature Vol. 469 Pg. 483 – 490

Krügel H, Solovjeva N. J, Fiedler G, Avraleva I.W and Schumann G. 1987. Occurrence of Phage-Like Particles and Growth Inhibition in *Streptomyces chrysomallus*. Journal of Basic Microbiology Vol. 27 Pg. 99 – 105

Kuchta R.D, Ilsley D, Kravig K.D, Schubert S and Harris B. 1992 Inhibition of DNA Primase and Polymerase Alpha by Arabinofuranosylnucleoside Triphosphates and Related Compounds. Biochemistry Vol. 31 Pg. 4720 – 4728

Kuchta R.D and Stengel G. 2010. Mechanism and Evolution of DNA primases. Biochimica et Biophysica Acta Vol. 1804 Pg. 1180 – 1189

Kusakabe T and Richardson C.C. 1996. The Role of the Zinc Motif in Sequence Recognition by DNA Primases. Journal of Biological Chemistry Vol. 271 Pg. 19563 – 19570

Kutateladze M and Adamia R. 2010. Bacteriophages as Potential New Therapeutics to Replace or Supplement Antibiotics. Trends in Biotechnology Vol. 28 Pg 591 – 595

LoBue P.A, Enarson D.A and Thoen T.C. 2010. *Tuberculosis* in Humans and its Epidemiology, Diagnosis and Treatment in the United States. The International Journal of Tuberculosis and Lung Disease Vol. 14 No. 10 Pg. 1226 – 1232

Lui J, Rosenberg E.Y and Nikaido H. 1995 Fluidity of the Lipid Domain of Cell Wall from *Mycobacterium chelonae*. Proceedings of the National Academy of Science Vol. 24 No. 92 Pg. 11254 – 11258

Lui J, Dehbi M, Moexk G, Arhin F, Bauda P, Bergeron D, Callejo M, Ferretti V, Ha N, Kwan T, McCarty J, Sirkumar R, Williams D, Wu J.J, Gros P, Pelletier J, DuBow M. 2004. Antimicrobial Drug Discovery through Bacteriophage Genomics. Nature Biotechnology Vol. 22 Pg 185 – 191

Marchand I, Nicholson A.W and Dreyfus, M. 2001. Bacteriophage T7 Protein Kinase Phosphorylates RNase E and Stabilizes mRNAs Synthesized by T7 RNA Polymerase. Molecular Microbiology Vol. 42 Pg. 767 – 776

Matthews I.I, Deacon A.M, Canaves J.M, McMullan D, Lesley S.A, Agarwalla S and Kuhn P. 2003. Functional Analysis of Substrate and Cofactor Complex Structures of a Thymidylate Synthase-Complementing Protein. Structure Vol. 11 Pg. 677 – 690

Mecsas J. 2002. Use of Signature-Tagged Mutagenesis in Pathogenesis studies. Current Opinion in Microbiology Vol. 5 Pg. 33 – 37

Mendelman L.V, Beauchamp B.B and Richardson C.C. 1994. Requirement for a Zinc Motif for Template Recognition by the Bacteriophage T7 Primase. EMBO Journal Vol.13 Pg. 3909–3916

Michel-Briand Y and Baysse C. 2002. The Pyocins of *Pseudomonas aeruginosa*. Biochimie Vol. 84 Pg. 449 – 510

Miller E.S, Kutter E, Mosig G, Arisaka F, Kunisawa T and Rüger W. 2003. Bacteriophage T4 Genome. Microbiology and Molecular Biology Review Vol. 67 Pg. 86 – 156

Mizuuchi K, Kemper B, Hays J and Weisberg R.A.1982. T4 Endonuclease VII Cleaves Holliday Structures. Cell Vol. 29 Pg. 357 – 365.

Mohl D.A and Gober J.W. 1997. Cell Cycle-Dependent Polar Localization of Chromosome Partitioning Proteins in *Caulobacter crescentus*. Cell 88, 675 – 684.

Moser M.J, Holley W.R, Chatterjee A and Main S. 1997. The Proofreading Domain of *Escherichia coli* DNA Polymerase I and other DNA and/or RNA Exonuclease Domains. Nucleic Acid Research Vol. 25 Pg. 5110 – 5118

Morris R.P, Nguyen L, Gatfield J, Viscont K, Nguyen K, Schnappinger D, Ehrt S, Liu Y, Heifets L, Pieters J, Schoolnik G and Thompson C.J. 2005. Ancestral Antibiotic Resistance in *Mycobacterium tuberculosis*. Proceedings of the National Academy of Science of the United States of America Vol. 102 Pg. 12200 – 12205

Mosig G. 1998. Recombination and recombination-dependent DNA replication in bacteriophage T4. Annual Review of Genetics Vol. 32 Pg. 379 – 413

Muňoz J, Extremera A.L, Arias J.M and Montoya E. 1987. Induction, Purification and Some Properties of Phage Tail-Like Particles from *Myxococcus coralloides* D. Microbios Vol. 49 Pg. 161 – 169

Myllykallio H, Lipowski G, Leduc D, Filee J, Forterre P and Liebl U. 2002. An Alternative Flavin-Dependent Mechanism for Thymidylate Synthesis. Science Vol. 297 Pg. 105 – 107

Myllykallio H, Leduc D, Filee J. and Liebl U. 2003. Life Without Dihydrofolate Reductase FolA. Trends in Microbiology Vol. 11 Pg. 220 – 223

Nakayama K, Takashima K, Ishihara H, Shinomyia T, Kageyama M, Kanaya S, Ohnishi M, Murata T, Mori H and Hayashi T. 2000. The R-Type Pyocin of *Pseudomonas aeruginosa* is Related to P2 Phage, and the F-Type is Related to Lambda Phage. Molecular Microbiology Vol. 38 Pg. 213 – 231

Nguyen L. 2005. FbpA-Dependent Biosynthesis of Trehalose Dimycolate id Required for the Intrinsic Multidrug Resistence, Cell Wall Structure, and Colonial Mophology of *Mycobacterium Smegmatis*. Journal of Bacteriology Vol. 187 Pg. 6603 - 6611

Nguyen L and Thompson C.J. 2006 Foundations of Antibiotic Resistance in Bacterial Physiology: The Mycobacterial Paradigm. TRENDS in Microbiology Vol. 14 No. 7 Pg. 304 – 312

Pedulla M.L, Ford M.E, Houtz J.M Karthikeyan T, Wadsworth C, Lewis J.A, Jacobs-Sera D, Falbo J, Gross J, Pannunzio N.R, Brucker W, Kumar V, Kandasamy J, Keenan L, Bardarov S. Kriakov J, Lawrence J.G, Jacobs W.R, Hendrix R.W and Hatfull G.F. 2003. Origins of Highly Mosaic Mycobacteriophage Genomes. Cell Vol. 113 Pg. 171 - 182

Philalay J.S. Palermo C.O, Hauge K.A, Rustad T.R and Cangelosi G.A. 2004. Genes Required for Intrinsic Multidrug Resistance in *Mycobacterium avium*. Antimicrobial Agents and Chemotherapy Vol. 48 Pg. 3412 – 3418

Phillips C.A. Spriggs T, Hedblom E, Cotton M.D, Utterback T.R, Hanna M.C, Nguyen D.T, Saudek D.M, Brandon R.C, Fine L.D, Firtchman J.L, Fuhrmann J.L, Geoghagen N.S.M, Gnehm C.L, McDonald L.A, Small K.V, Fraser C.M, Smith H.O and Venter C. 1995. Whole Genome Random Sequencing and Assembly of Haemophilus influenzae Rd. Science Vol. 269 Pg. 496 – 512

Pope W.H, Jacobs-Sera D, Russell D.A, Peebles C.L, Al-Atrache Z, Alcoser T.A, Alexander L.M, Alfano M.B, Alford S.T, Amy N.E, Anderson MD, *et al*. 2011
Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. PLoS One Vol. 5 e16329

Powell I.B, Tulloch D.L, Hillier A.J and Davidson B.E. 1992. Phage DNA synthesis and host DNA degradation in the life cycle of *Lactococcus lactis* bacteriophage c6A. Journal of General Microbiology Vol. 138 Pg. 945 – 950

Purwantini E. Gillis T. P and Daniels L. 1997. Presence of F420-Dependent Glucose-6-Phosphate Dehydrogenase in Mycobacterium and Nocardia Species, but Absence from Streptomyces and Corynebacterium Species and Methanogenic Archaea. FEMS Microbiology Letters Vol.146 Pg. 129 – 134

Raaijmakers H, Vix O, Törõ I, Golz S. Kemper B and Suck D. 1999. X-Ray Structure of T4 Endonuclease VII: a DNA Junction Resolvase with a Novel Fold and Unusual Domain-Swapped Dimer Architecture. The European Molecular Biology Organization Journal Vol. 18 Pg. 1447 – 1458

Reha-Krantz L.J and Nonay R.L. 1993. Genetic and Biochemical Studies of Bacteriophage T4 DNA Polymerase 3'-5'-Exonuclease Activity. Journal of Biological Chemistry Vol. 268 Pg. 27100 – 87108

Reichard P. 1993 From RNA to DNA, Why so Many Ribonucleotide Reductases? Science Vol. 260 Pg. 1773 – 1777

Rifat D, Wright N.T, Varney K.M, Weber D.J, and Black L.W. 2008 Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target. Journal of Molecular Biology Vol. 375 Pg. 720 – 734

Roucourt B and Lavigne R. 2009. The Role of Interactions between Phage and Bacterial Proteins within the Infected Cell: A Diverse and Puzzling Interactome. Environmental Microbiology Vol. 11 Pg. 2789 – 2805

Rybniker J, Kranne S and Small Pl. 2006. Host Range of 14 Mycobacteriophages in *Mycobacterium ulcerans* and seven other mycobacteria including *Mycobacterium tuberculosis* – Application for Identification and Susceptibility Testing. Journal of Medical Microbiology Vol.55 Pg 37 – 42

Sampathkumar P, Turley S, Ulmer J.E, Ho G.H, Sibley C.H and Hol W.G.J. 2005. Structure of the Mycobacterium tuberculosis Flavin Dependent Thymidylate Synthase (MtbThyX) at 2.0 Å Resolution. Journal of Molecular Biology Vol. 352 Pg. 1191 – 1104

Sharpe M.E and Errington J. 1997. The *Bacillus subtilis soj-spo0J* Locus is Required for a Centromere-Like Function Involved in Prespore Chromosome Partitioning. Molecular Microbiology Vol. 21 Pg. 501 – 509.

Sharples G.J. 2001. The X Philes: Structure-Specific Endonucleases that Resolve Holliday Junctions. Molecular Microbiology. Vol. 39 Pg. 823 – 834

Shibayama Y. 2006. Rapid Screening for Antimicrobial Genes in Novel Nocaridophages. MSc Dissertation. University of the Witwatersrand, Johannesburg

Shibayama Y and Dabbs E.R. 2011. Phage as a Source of Antibacterial Genes: Multiple Inhibitors Encoded by Rhodococcus Phage YF1. Bacteriophage Vol. 1 Pg. 195 – 197

Skurnik M and Strauch E. 2006. Phage Therapy: Facts and Fiction. International Journal of Medical Microbiology Vol. 296 Pg. 5 – 14

Solaro P.C, Birkenkamp K, Pfeiffer P and Kemper B. 1993 Endonuclease VII of Phage T4 Triggers Mismatch Correction in Vitro. Journal of Molecular Biology Vol. 230 Pg. 868 – 877

Strauch E, Schaudinn C, 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. Infection and Immunity Vol. 72 Pg. 7030 – 7036

Suttle C.A, Chan A.M, Cottrell M.T. 1990. Infection of Phytoplankton by Viruses and Reduction of Primary Productivity. Nature Vol. 347 Pg. 467 – 469

Suttle C.A 2007 Marine Viruses – Major Players in the Global Ecosystem. Nature Reviews Microbiology Vol. 5 Pg. 801 – 812

Takahashi I. and Marmur J. 1963 Replacement of Thymidylic Acid by Deoxyuridylic Acid in the Deoxyribonucleic Acid of a Transducing Phage for *Bacillus subtilis*. Nature Vol. 197 Pg. 794 – 795

Tock M.R and Dryden D.T.F. 2005. The Biology of Restriction and Anti-Restriction. Current Opinion in Microbiology Vol. 8 Pg. 466 – 472

Tokunaga T, Kataoka T and Suga K. 1970. Phage Inactivation by an Ethanol-Ether Extract of *Mycobacterium smegmatis*. American Review of Respiratory Disease Vol. 101 Pg. 309 – 313

Torella F and Morita R.Y. 1979 Evidence by Electron Micrographs for a High Incidence of Bacteriophage Particles in the Waters of Yaqunia Bay, Oregon: Ecological and Taxonomical Implications. Applied and Environmental Microbiology Vol. 37
Pg. 774 – 778

Wang Z. and Mosbaugh D.W. 1989 Uracil-DNA Glycosylase Inhibitor Gene of Bacteriophage PBS2 Encodes a Binding Protein Specific for Uracil-DNA Glycosylase. Journal of Biological Chemistry Vol. 264 Pg. 1163 – 1171

Warren R.A.J. 1980. Modified Bases in Bacteriophage DNAs. Annual Review in Microbiology Vol. 34 Pg.137 – 158

Wilson E.M and Trauner D. 2007. Concise Synthesis of the Bacterial DNA Primase Inhibitor (+)-Sch642305. Organic Letters Vol. 9 Pg. 1327 – 1329

Wilson W.H, Fuller N.J, Joint I.R and Mann N.H. 1999. Analysis of Cyanophage Diversity and Population Structure in a South-North Transect of the Atlantic Ocean. Bulletin de l'Institut Oceanographique (Monaco) Vol. 19 Pg. 209 – 216

WHO. 2009. Global *Tuberculosis* Control: Short Update to 2009 Report. World Health Organization

WHO. 2010. The Global Plan to Stop TB 2011 – 2015: Transforming the Fight Towards Elimination of *Tuberculosis*. World Health Organization

Yoong P. 2004. Identification of a Broadly Active Phage Lytic Enzyme with Lethal Activity against Antibiotic-Resistant *Enterococcus faecalis* and *Enterococcis faecium*. Journal of Bacteriology Vol. 186 Pg. 4808 – 4812

Young R. 1992 Bacteriophage Lysis: Mechanism and Regulation. Microbiological Review Vol. 56 Pg. 430 – 481

Young P. Őhman M and Sjöberg B.M. 1994. Bacteriophage T4 Gene 55.9 Encodes an Activity Required for Anaerobic Ribonucleotide Reduction. The Journal of Biological Chemistry Vol. 269 Pg. 27815 – 27818

Zignol M, Hossenin S, Wright A, Lambregts-van Weezenbeek C, Nunn P, Watt C.J, Williams B.G and Dye C. 2006. Global Incidences of Multidrug Resistant *Tuberculosis*. The Journal of Infectious Disease Vol. 194 Pg. 479 – 485